

**INTRODUCTION TO**

**EMPIRICAL AND SCIENTIFIC**

**PSYCHOLOGY**

**DLBPSEEWP01\_E**



# LEARNING OBJECTIVES

***2024-06-12 11:33:03***

--------------------------------------------

Not true

The course book **Introduction to Empirical and Scientific Psychology** is designed to provide an overview of psychology as a scientific discipline. It focuses on the empirical methods of obtaining and improving scientific knowledge. You will be given the opportu- nity to understand and implement scientific methods as well as gain knowledge and nec- essary skills to critically evaluate research-related topics that are relevant to psychology.

Basic scientific terms will be explained, and you will learn to distinguish between the pro- fessional and linguistic use of terms that are essential for research psychological commu- nication. For example, you will be introduced to terms such as variables, hypotheses, operationalization, and measurement.

In this course, you will also learn to apply research and scientific knowledge in designing a research project, what it involves—from planning to addressing ethical dilemmas and decisions to the implementation and presentation of research findings. The course book will furthermore introduce you to quantitative, qualitative, and mixed research method approaches with a focus on quantitative research methods. In addition, it will explain why psychology is an empirical science. Overall, the course book will offer you a robust sum- mary of scientific psychology that helps you acquire the basic knowledge to recognize and evaluate methodological content relevant to statistics, research methods, and diagnostics at later stages of their studies.

# UNIT 1

## PSYCHOLOGY AS AN EMPIRICAL SCIENCE

**STUDY GOALS**

On completion of this unit, you will be able to ...

* explain why psychology is an empirical science.
* distinguish between scientific and non-scientific approaches.
* describe the available psychological research methods.
* understand and formulate scientific hypotheses.
* identify and describe the different types of available variables we often come across in research.

### 1. PSYCHOLOGY AS AN EMPIRICAL SCIENCE

#### Introduction

The primary motivation for most people to study psychology is the desire and interest to understand human behavior and consequently to help and support the people who need psychological and psychiatric attention. For that reason, it often comes as a surprise to many new psychology undergraduates that a substantial number of courses taught in a psychology degree deal with empirical science, methodology, and experimentation as well as statistics. A major goal of every psychology degree is to help students understand, evaluate, and produce scientific findings independently. These are paramount skills to have as a qualified psychologist, regardless of the field of work. For example, skills rele- vant to scientific methods and statistics are essential to be able to work as a clinical psy- chologist who practices psychotherapy so that one can decide on the appropriate, evi- dence-based interventions or psychotherapeutic approaches. In human resources, psychologists need the same type of expertise and skills to conduct employee and employer surveys. School psychologists should also have and apply necessary research and scientific skills to select appropriate procedures in cases where they need to decide on different learning support options for students. These are just a few examples demon- strating how important it is for psychologists in any field to understand psychology as an empirical science and to develop a scientific way of thinking. This unit explains why psy- chology is a scientific discipline and gives an overview of relevant scientific principles and methods.

#### Subject of Psychology

If we were to give a definition for psychology, we would describe it as a scientific field that deals with human behavior and actions as well as mental processes. Although this might seem an overly simplistic definition at first, it includes fundamental questions regarding emotion and feelings, such as parent-child interaction, the experience of human well- being, and how to improve people’s experience of depression. An interesting example that is embodied in the definition of psychology provided above relates to the existence of stereotypes in society. People learn about differences in behavior between men and women from an early age, and a widespread belief to this end suggests that women talk more than men and that this difference is even biologically driven (Cuttler et al., 2019). But is this true?

A group of psychologists decided to investigate this claim and conducted a research study (Mehl et al., 2007). They were driven by the fact that, despite the widely accepted belief that women are more talkative than men, no one had counted the words spoken daily by them. In their study, Mehl et al. (2007) gave audio recorders to male and female college students, and asked them to continue their daily routine as usual. The findings have shown that the number of words spoken daily by men and women did not differ. Hence, the research team concluded that there is no scientific basis for such a stereotype (Mehl

et al., 2007). Considering this example, one could conclude that psychology not only wishes to understand and provide an explanation for human experience, behavior, and actions, but also describes, changes, and even predicts human behavior (Allport, 1940; Cuttler et al., 2019).

##### Description

Psychological characteristics are classified, named, defined, and recorded. In clinical psy- chology, psychiatric disorders, such as anxiety, depression, addiction, or eating disorders, are classified into categories. Furthermore, clinical psychology examines whether there is an association between gender and certain disorders. A representative German study, for example, found that it is more common for girls rather than boys between the ages of 11 and 19 to suffer from depressive symptoms (Ravens-Sieberer et al., 2007). A literature review published by Culbert et al. (2021) also concludes that females might be more prone to eating disorders than males.

##### Explanation

It is important to discover which trait might cause a specific psychological phenomenon or problem. The two previous examples support the idea of an association between gen- der and depression or eating disorders. The reasons behind this are not entirely known yet, but possible explanations point to biological factors such as hormonal differences and neurobiological mechanisms that potentially manifest differently in men and women.

##### Prediction

Experiences and behaviors can be predicted. For example, it is possible to find out which adolescents are more likely to have a worsening of their symptoms. Based on statistical analysis we can predict that alcohol use disorder patients with smaller orbitofrontal cortex volume have increased chances of relapsing 6 months following their treatment. Follow- ing such predictions, which are the result of necessary statistical analysis, clinicians can implement appropriate psychotherapeutic interventions (Zois et al., 2017).

##### Change

Psychological characteristics can often be influenced by several factors and are therefore subject to constant change. For instance, a therapeutic intervention was carried out for the young people from the example above to reduce their depressive symptoms.

##### Key Elements of the Scientific Approach

Psychology follows the scientific approach, which is characterized by three main ele- ments: systematic empiricism, empirical questions, and public knowledge (Stanovich, 2010).

###### Systematic empiricism

***2024-06-12 11:42:40***

--------------------------------------------

i.e. and e.g. should always be in parentheses

Systematic empiricism elates to observations of the world in a methodical way (planning, implementing, recording, analyzing), and allows scientists to express their ideas and intu- itions combined with their systematic observations (Cuttler et al., 2019). Empirical ques- tions are connected to systematic empiricism because they concern matters about how the world is and can be answered only through systematic observations.

###### Empirical questions

Going back to the example about gender talkativeness (Mehl et al., 2007), the question would be whether women are more talkative than men, and this can be answered by implementing a systematic observation, i.e., recording the amount of words women and men speak. It is important to consider that not all questions about the world can be sys- tematically tested and observed through the scientific lens (Cuttler et al., 2019). Examples are questions of taste and personal preference, i.e., how beautiful or ugly, or how good or bad something is.

###### Public knowledge

After scientists have defined their empirical questions, put them into systematic observa- tion, determined what is true and what is not (e.g., that women are not more talkative than men), they need to make this knowledge accessible. This can be achieved by writing and publishing a research article in academic journals (like those you are expected to read throughout this course). In this way, a scientist has the chance to explain the aims of their research compared to previous relevant research, give a detailed overview of their meth- odological approach, and finally show their statistical analysis and conclusions (Cuttler et al., 2019).

In psychology, research questions are examined using empirical methods. Empirical research methods are based on concrete practice and knowledge. In empirical sciences such as psychology, physics, biology, or mechanical engineering, hypotheses are formu- lated (emerging from research questions) and challenged by reality (what is happening vs. what was believed). The scientific approach observes and measures facts and conclusions that are not drawn by logical, purely intellectual assumptions and beliefs, as in the case of philosophy. The example of gender talkativeness illustrates the advantage of empirical research over logical or intuitive reasoning. In psychology, phenomena that cannot be observed directly, e.g., intelligence, and attitudes, are often examined and need to be determined first. Therefore, in addition to the observation methods used in all empirical sciences, psychology employs various methods of empirical social research, such as inter- views or surveys. In psychology, there is a critical debate on the extent to which empirical research results correspond to reality.

##### Consequences of Lack of Scientific Knowledge: An Example

Psychologists often hold positions of influence, which means their opinion matters to cer- tain groups of people as they are perceived as authority figures. The professional guide- lines strongly encourage psychologists to strive for a high level of competence and to be

aware of their responsibilities for clients. It is possible, however, that psychologists divert from their profession’s official guidelines and end up giving their clients information that has little scientific value. The consequences can be serious. For example, autism is often diagnosed by trained psychotherapists in a psychotherapeutic setting. Children who have autism often show limited social competence, as well as emotional and communication difficulties. In the 90s, a non-scientific article which was eventually retracted depicted autism as a result of the measles vaccination (Davidson, 2017). If a psychologist shared those views because of their lack of scientific knowledge on this matter and presented such information to parents with autistic children, negative consequences would arise. The parents would feel guilty for getting their children vaccinated or might prevent their children from getting vaccinated in the future, and of course spread the (false) information about the dangers of vaccinating to other parents.

Scenarios like this happened in reality—especially in Great Britain in the 90s, because of findings showing that the measles vaccine was associated with the diagnosis of autism in children. These findings were later withdrawn after they had been proven to be untrue (Davidson, 2017). Regardless, the rate of measles vaccinations had fallen from 92% in 1996 to 85% in 2006, correspondingly, the number of measles cases increased (McIntyre & Leask, 2008). A trained psychologist should always act responsibly and according to the professional guidelines (DFG, 2022). They have to base their professional opinion on relia- ble scientific knowledge. A psychologist should also avoid expressing an opinion on mat- ters outside their professional domain, knowledge, and training. Vaccination concerns the medical field, and a responsible psychologist should suggest their clients to discuss such issues with trained medical professionals.

#### How Is Knowledge Created?

Science aims at generating new knowledge. Psychology is no different in that respect; it seeks answers to questions related to human experience, behavior, and action. Such questions are often relevant to our everyday life experience, i.e., when a teacher wishes to know whether boys are more aggressive than girls. This is an example of a relevant ques- tion for psychological research: Are there gender differences in aggressive behavior in adolescents? It is not the questions themselves that distinguish scientific psychology from folk psychology, but the approach one takes in answering such questions. Going back to the example with the teacher, they might assume boys to be more aggressive than girls. A teacher’s opinion, however, is not scientifically justified. According to Björkqvist (2018), it is false and ridiculous to claim that aggression is gender-dependent. This insight is derived from scientific proof based on research and published works on the topic (Björkqvist, 2018). Björkqvist’s work has shown that the frequency of aggressive behavior in girls and boys does not differ. The difference, however, lies in the form of aggression: Boys are more often physically aggressive, whereas girls rather use indirect verbal ways of expressing aggression. Another example which is particularly relevant for forensic psy- chology is false confession. People are certain that no one would ever confess to a crime they have not committed unless they were violently forced to do so. Research, however,

tells us otherwise: According to research findings, false confessions are very common, and they are made for various reasons which are described in the review by Kassin and Gud- jonsson (2004).

##### Ways of Knowledge Acquisition

How could the teacher conclude that boys are more prone to aggression without support- ing scientific evidence? What are the limitations of their conclusions based solely on intu- ition or observation? Obtaining knowledge can be accomplished through various means, each with its own set of advantages and disadvantages (Cuttler et al., 2019).

###### Intuition

We frequently rely on our intuition, also known as gut feeling or instinct, to navigate our behavior and gain new knowledge (Gigerenzer, 2008). However, this approach to acquiring knowledge is not grounded in rational thinking and is solely based on our instinctual response to situations (Cuttler et al., 2019). Thus, relying on intuition alone may not be a dependable guide in the quest to acquire knowledge or learn. It is frequently affected by cognitive or motivational biases rather than scientifically driven conclusions (Cuttler et al., 2019). However, weighing numerous alternatives and possibilities can be a laborious deci- sion-making process, so intuition may occasionally be the optimal choice (Gigerenzer, 2008; Cuttler et al., 2019).

###### Authority

It is a common occurrence for individuals to acquire new knowledge through authority fig- ures. Novel ideas are typically accepted solely based on the endorsement of authoritative figures. Some examples of authority figures include parents, government officials, media, healthcare professionals, religious leaders, and university professors (Cuttler et al., 2019). It is ideal for individuals to trust authority figures, but history has shown us that this is not always the case. Many of the worst crimes against humanity occurred when society failed to question authority, such as during World War II. However, this does not suggest that we should always distrust authorities in the pursuit of knowledge. Rather, we must learn to evaluate the credibility of information we are given and determine whether authorities have a potential incentive to mislead the public (Cuttler et al., 2019).

###### Rationalism

Acquiring new knowledge frequently involves rationalism and the use of logic and reason- ing. Conclusions are drawn by adhering to logical rules and methods, even if observation is absent. Some argue that this is a reliable method of gaining knowledge, as long as logi- cal rules are followed without errors (Cuttler et al., 2019). For instance, if presented with the principle that all swans are white and the observation of a swan, one could logically conclude that this particular swan is white without further observation. However, this approach can be problematic if the initial principle is incorrect. Not all swans are white; there are also black swans (Cuttler et al., 2019).

###### Empiricism

Knowledge acquisition stems from experience and observation. While observation lies at the heart of the scientific method, subjective evaluations must be set aside unless explic- itly marked as such. Overreliance on observation and experience can lead to erroneous conclusions, as evidenced by the belief that all swans are white. Similarly, the perception of a flat earth is not an accurate representation of reality. Empiricism, while valuable, is not the sole source of gaining new knowledge and should be approached with caution. Nevertheless, the scientific method heavily depends on empiricism provided that it is structured and systematic (Cuttler et al., 2019).

##### The Scientific Method

This approach integrates intuition, authoritative sources, rationalism, and empiricism to gather information and evidence, analyze novel concepts, and address lingering inquiries in order to advance knowledge (Cuttler et al., 2019). It is crucial to note that for testing such innovative ideas or obtaining fresh information and evidence, a systematic approach is employed, with observations conducted meticulously under strictly controlled condi- tions (Cuttler et al., 2019). The scientific method is generally considered the most reliable approach for generating new knowledge, but it is not without limitations. Practical factors such as time and resource constraints may limit the application of the scientific method, and some questions may not lend themselves to this approach (Cuttler et al., 2019).

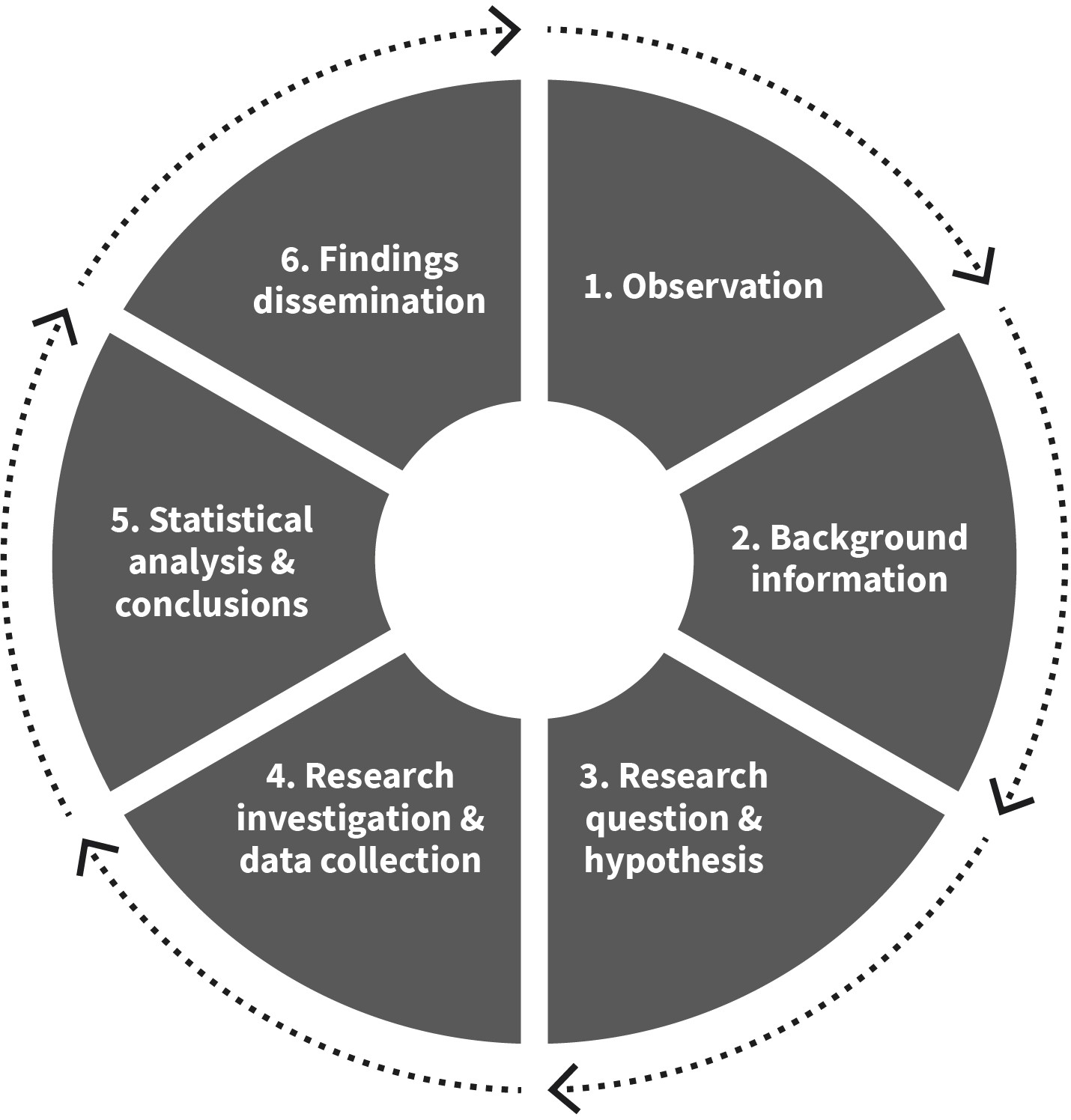
The scientific method is a stepwise process for resolving questions through testing. To begin, a researcher makes an observation of a phenomenon of interest or poses a ques- tion requiring an answer. For instance, a clinical psychologist might observe that only younger men attend therapy in the addiction clinic and wish to quantify this observation scientifically. Background information collection is a necessary precursor to data. For instance, a review relevant published records on age and addiction treatment before craft- ing a hypothesis to explain the observation or research question. If a clinical psychologist can demonstrate that young individuals with addiction attend therapy more frequently than older patients, this becomes a scientifically established fact. Based on the hypothe- sis, the researcher will generate a testable prediction that can be measured through the application of empirical analysis, so that clinical psychologists or researchers can attain insight into whether statistical findings align with the hypothesis. For instance, a clinical psychologist may want to confirm the observation that younger patients favor addiction therapy more than older ones. Does the data analysis confirm this or differ from the psy- chologist’s observations? The documentation and dissemination of findings to the scien- tific community constitute the final step in the process.

Figure 1: The Scientific Method

***2024-06-12 11:37:58***

--------------------------------------------

Verify how true this is around the world



Source: Evangelos Zois (2022).

##### Good Scientific Practice

Knowledge acquisition through empirical research and science requires certain standards and criteria. To meet these criteria, research must be conducted in a trustworthy and integrity-driven manner. In pursuit of new knowledge, psychologists must rely on theoreti- cal considerations, past empirical findings, and reliable research methods to draw their conclusions. Researchers adhere to a rigorous process that is both systematic and tracea- ble, with detailed documentation, rigorous evaluation, and monitoring, facilitating another undertaking of the research. For a thorough overview and further elaboration on Good Scientific Practice, consult the Code of Good Scientific Practice (GSP) established by the German Research Society (DFG, 2022). All scientific and academic institutions are obli- gated to comply with the GSP code and ensure that their employees and students are equipped with the required resources and knowledge to adhere to the code as well (DFG, 2022). In the pursuit of acquiring scientific knowledge, it is essential to strictly adhere to some fundamental guidelines and regulations which are outlined below:

###### Research question

Empirically verifiable and clearly formulated research questions are essential in the pur- suit of knowledge. These questions serve as the foundation of scientific inquiry and must be concrete and falsifiable. It is imperative to avoid vague and overly general statements. A scientifically trained researcher seeks to answer research questions using appropriate and sound methods. In situations where novel assessment techniques are required, researchers typically concentrate on implementing the required quality control measures while adhering to a predefined set of implementation guidelines (DFG, 2022).

###### Research status

The consideration of the current state of research is crucial. Scientists should be aware of existing theories and previous empirical findings when formulating research questions and selecting corresponding research methods. This is essential for studying novel ques- tions with no, inadequate, or conflicting results. Researchers often formulate research questions to replicate and validate prior findings and results (Cuttler et al., 2019).

###### Data

The systematic collection, processing, and analysis of data should adhere to established scientific methods. Empirical data should be gathered and examined using appropriate research designs, whenever possible, with a representative sample and validated data col- lection instruments and strategies. It is crucial to adhere to scientific data quality criteria. For instance, standardized questionnaires are regarded as an objective method of data collection.

###### Ethics

In the field of psychology, there are sets of important principles, professional guidelines, research ethics, and codes of conduct formulated by organizations such as the Federal Association of German Psychologists (BDP) and the German Society for Psychology (DGPs) for psychologists in Germany, and the American Psychological Association (APA) in the United States, as well as similar associations worldwide. These guidelines, particularly in the USA and Europe, have significant overlap. Ethics are of utmost importance as they involve principles and guidelines that have been carefully developed to protect individu- als who participate in research studies, commonly referred to as research participants. For example, these guidelines suggest that research participants must be fully informed about the objectives of the research project they are considering participating in before provid- ing their consent. Furthermore, participants should have the freedom to withdraw from the project at any point, even if they had previously given written consent for their partici- pation (DFG, 2022; Cuttler et al., 2019).

###### Transparency

The research process must always be closely monitored, transparent, and thoroughly documented. The data should be archived in a manner that allows for traceability and ver- ification at any point during and after the data collection process. It is essential to docu-

ment and describe a research project in sufficient detail to provide all the necessary infor- mation for potential replication. For instance, if a researcher adheres to this principle, another researcher should be able to replicate the same project by studying the written procedure provided by the original researcher. Collected data should be kept anonymous, particularly in electronic form, for a period of several years (depending on the nature of the research, ranging from five to ten years). Additionally, published data should be made available to readers and researchers in an anonymous manner, enabling them to work with and reanalyze the data for the purpose of verifying the original analysis and findings (DFG, 2022).

###### Results

The results or findings obtained from a research project should be interpreted with two key considerations in mind. Firstly, it is important to reflect on the potential limitations that may exist within the project. Secondly, previous findings should be considered during the interpretation process. The results should be understood within the context of the original research questions, expectations, and the broader research context. It is advisable to avoid generalizing standalone results. In the field of psychology, findings are not regarded as proven facts in the same way as they would be in mathematics. Instead, results are interpreted as indications of certain established facts (Cuttler et al., 2019).

###### Dissemination of findings

Results and findings should be made publicly accessible, through publication in reputable peer-reviewed scientific journals. Peer review is a crucial process that every research arti- cle undergoes before publication. During this process, other experts or scientists provide their expert opinion to the editor of the scientific journal regarding the scientific paper submitted for publication. The peer review process is blind, meaning that the identities of the researcher and the reviewers are not disclosed to each other to maintain objectivity. Reviewers provide feedback on the methods, analysis, data collection strategies, and other aspects of the work, with the overall goal of processing and improving the submitted work. Based on the peer review process, the editor will decide on whether to publish the paper or not.

Traditionally, published work was only accessible to academics in universities and research centers. However, open science sources are becoming increasingly popular, mak- ing scientific knowledge more publicly available. Other ways to disseminate research find- ings include presenting them in lecture theaters at universities in front of a specialized audience. Additionally, lectures open to the public or individuals involved in policy-mak- ing can also be used as a means to share research findings (Cuttler et al., 2019).

##### Philosophy of Science

The philosophy of science is a branch of philosophy that originally deals with the empiri- cal processes, methods, and implications of science. It is concerned with assumptions about the nature of objects in the natural world under investigation, the reliability of theo- ries, the limits and prerequisites of research methods, as well as the ultimate purpose and responsibility of science in society (Rosenberg, 2005). Additionally, the philosophy of sci-

ence describes the extent to which research methods can be used to verify reality by com- bining theory and empirical data. For the purposes of this module, two ways of scientific knowledge acquisition driven by theory and empirical data will be described, specifically the inductive and deductive research approaches (Streefkerk, 2022). Their main difference lies in the approach: either the development of a theory (inductive), or the testing of an existing theory (deductive) (Cuttler et al., 2019; Streefkerk, 2022). Deductive reasoning moves from general principles to specific observations, whereas inductive reasoning pro- gresses in the opposite direction (Streefkerk, 2022). Both methods have advantages and disadvantages and are frequently utilized in research, either in conjunction or individually (Cuttler et al., 2019; Streefkerk, 2022).

###### Inductive method

In reality, research often involves investigating subjects to gain a better understanding of them, but but this can become challenging due to a lack of published literature. In such cases, it is common to follow the inductive approach to gather more information. This involves starting with observations, examining them to identify patterns, and finally using the data obtained from these observations to either develop a theory or draw broad con- clusions (Streefkerk, 2022). The main concern regarding the inductive approach is that its conclusions cannot be verified, but rather can only be invalidated (Streefkerk, 2022). An example by Streefkerk (2022) shows that if the aim is to observe the flight schedule of sev- eral budget airlines, suppose 1,000 in number, and each of them is delayed, the new theory states that budget airlines have delays. However, this theory cannot confirm that flight 1,001 will also be delayed. A greater number of flights would be required for obser- vations. As will be discussed later, a larger sample size would result in a more reliable con- clusion (Streefkerk, 2022).

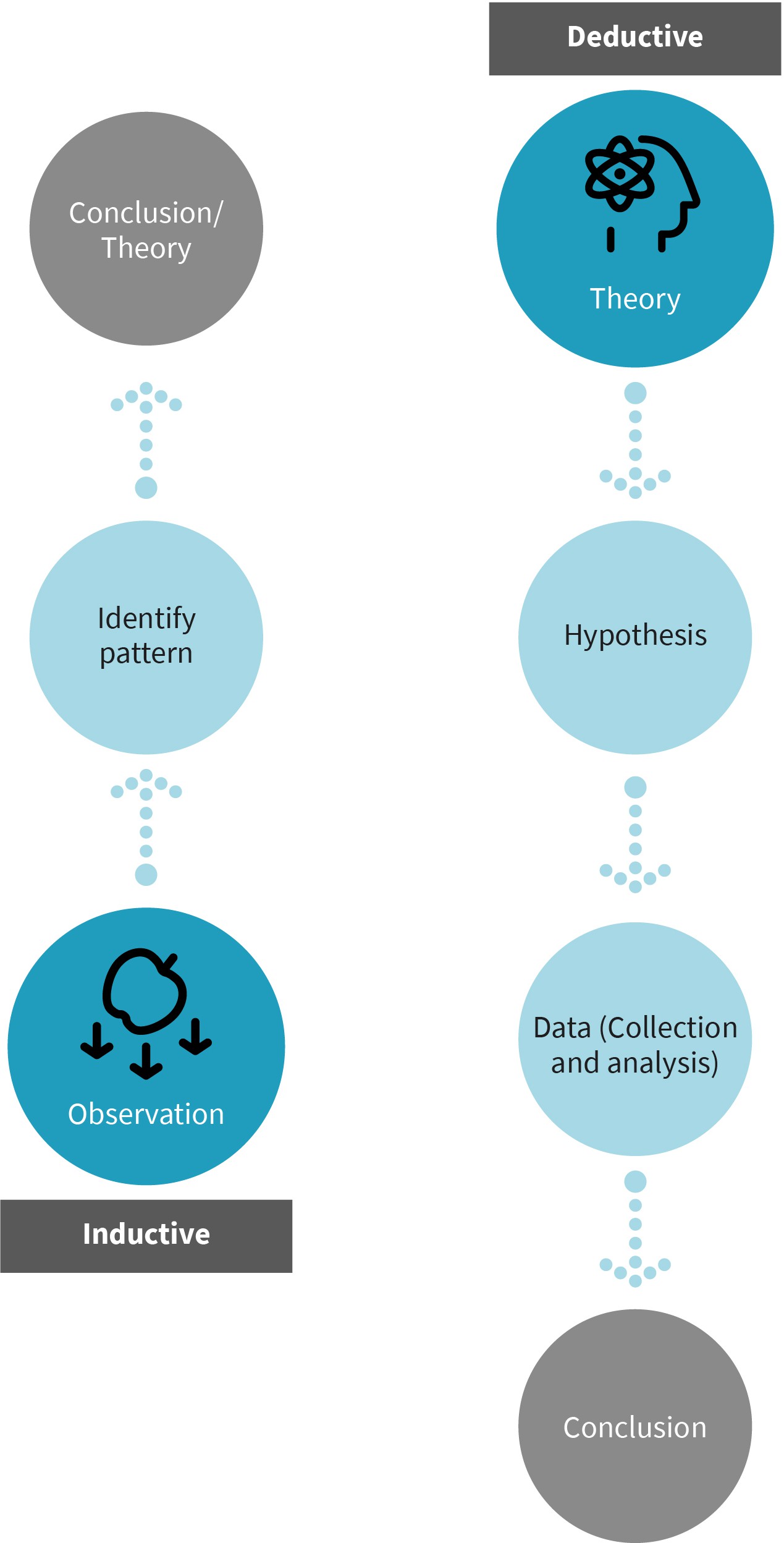
###### Deductive method

In this approach, the result of the inductive approach, i.e., a theory, is tested. Firstly, an existing theory such as “budget airlines always have delays” is stated (Streefkerk, 2022). A **falsifiable** hypothesis originating from the existing theory is formulated, e.g., “customers travelling with budget airlines experience delays” (Streefkerk, 2022). Information (data) is gathered and analyzed after that. Finally, the original hypothesis is verified or falsified, based on the statistical analysis. An example for such a result would be: “Five out of 100 flights were delayed.” A disadvantage of the deductive approach is its complete reliance on the theory derived from inductive reasoning, assuming the theory’s validity (Streef- kerk, 2022).

**Falsifiable**

scientific claims conveyed in such a way that obser- vations (if made) would hold as evidence against them

Figure 2: The Process of the Inductive and Deductive Research Approach



Source: Evangelos Zois (2022), based on Streefkerk (2022).

#### Systematics of Scientific Methods

In this section, the different types of research methods relevant to applied psychology (diagnostics, psychological interventions, evaluation) will be discussed. Descriptions and distinctions between quantitative and qualitative methods, as well as mixed methods, will be provided. It is important to note the specific criteria for data collection and analysis.

##### Methods of Data Collection and Analysis

Data collection refers to the procedures researchers employ to record, store, and analyze human experiences, behaviors, and actions. For instance, if the objective is to gain a better understanding of human memory, researchers may ask participants a question such as, “Do you remember events before the age of three?” Participants are then required to respond by selecting one of three options provided: “never,” “sometimes,” or “often.” Each participant’s answer is subsequently stored, with each word replaced by a corresponding number. In this instance, the number 1 represents “never,” 2 represents “sometimes,” and 3 represents “often.” This example illustrates the most basic method of collecting and organizing data in a meaningful manner for empirical examination of the research ques- tion. To analyze this data, one could, for instance, calculate the frequency of participants who responded with “sometimes” or “often” (which equals a number greater than 1) in terms of the number of events they can recall before the age of three.

##### Quantitative Research Methods

Quantitative research methods involve the systematic collection and analysis of standar- dized data that is stored in numerical form. These methods focus on objective measure- ment and analysis using numerical data (McLeod, 2019). Numerical data in psychological research is typically obtained through the administration of various standardized tools and instruments, such as intelligence tests, reaction time tasks, and tools that record physiological responses like sweat. Standardized questionnaires are commonly used to collect data in psychological research. Regardless of the specific instrument or tool employed, the response options are precisely defined and can be consistently stored. The primary objective of quantitative research is to describe relationships between variables, predict outcomes, and, most importantly, generalize findings to the broader population (McLeod, 2019). Quantitative research methods involve collecting data from a large group of individuals with specific characteristics (known as a sample). These methods employ a deductive approach, drawing conclusions based on existing theories and empirical evi- dence.

Critics often argue that quantitative research methods are limited because data collection typically occurs in controlled environments, such as a laboratory or non-natural settings. In these settings, participants respond to questions or perform tests based on the specific research design. However, quantitative research fails to capture the individual reasons behind each participant’s choices, and there is limited opportunity to explore the personal meaning of each question for each participant separately (McLeod, 2019). Another criti- cism of quantitative research methods pertains to the potential lack of adequate statisti- cal knowledge. As mentioned earlier in this subsection, quantitative research involves

analyzing and interpreting recorded numerical data using statistical analysis. If a researcher lacks the necessary statistical knowledge, it can have negative consequences for the accuracy of the analysis, interpretation of findings, and ultimately the conclusions drawn from the research (Cuttler et al., 2019; McLeod, 2019). Furthermore, there is the issue of sample size in quantitative research. While a larger sample size is generally prefer- red, it is not always feasible to obtain one. Quantitative research studies often end up with relatively small samples, which are sometimes considered less reliable and may limit the ability to make generalizations about the broader population (McLeod, 2019)

##### Qualitative Research Methods

Qualitative research is frequently employed in psychological studies and is characterized by research questions that are broad and less specific. This approach is chosen because the primary objective of qualitative research is to gain a comprehensive understanding of the behavior of interest. In qualitative research, the focus is on exploring the reasons behind individuals’ responses, as inter-individual differences are considered relevant and valuable in understanding the phenomenon under investigation (Cuttler et al., 2019). In qualitative research methods, the emphasis is placed on the depth and richness of data collected from each participant, rather than the size of the overall sample. This is a key distinction between qualitative and quantitative research methods (McLeod, 2019). Quali- tative research often involves collecting non-standardized data through methods such as naturalistic observations or interviews, which are saved in their original form (e.g., as text rather than numerical data). The approach to qualitative research typically follows an inductive approach, where theories and hypotheses are generated based on the data col- lected. Qualitative investigations are particularly valuable when there is limited existing knowledge on a topic and new theories and hypotheses need to be developed (McLeod, 2019).

##### Mixed Research Methods

As discussed earlier, both qualitative and quantitative research methods have their own strengths and limitations. In contemporary psychological research, there is a preference for integrating both methodological approaches and employing a mixed methods research methodology. This approach allows researchers to benefit from the advantages of both qualitative and quantitative research designs. Depending on the research question and the specific topic being investigated, a mixed methods approach can be utilized to employ either inductive or deductive reasoning. By combining these approaches, researchers can gain a more comprehensive understanding of the phenomenon under study (McLeod, 2019).

##### Research Methods Relevant to Applied Psychology

In the field of applied psychology, several methods are commonly used, including diag- nostics, intervention, and evaluation.

###### Diagnostics

Psychological diagnostics is an empirically-based assessment method used to identify and understand the characteristics and differences of individuals or groups. Psychologists make observations and verify them using scientific methods such as tests, questionnaires, interviews, behavioral observations, and medical history. Similar individuals are catego- rized into groups to answer diagnostic questions and determine appropriate measures, such as therapy, training/coaching, or changes in current conditions. Diagnostics play a crucial role in various areas of applied psychology. For instance, in clinical psychology, diagnostics are used to make diagnoses, which guide professionals like clinical psycholo- gists or psychiatrists in planning the most effective therapeutic approach. In educational psychology, diagnostics are employed to assess students’ performance relative to their peers or to determine the suitability of specific educational methods. Both quantitative methods (tests, questionnaires) and qualitative methods (interviews, observations) are utilized in diagnostics.

###### Intervention

Diagnostics often lead to the development and implementation of appropriate psycholog- ical interventions, primarily in clinical and health psychology. The goal of interventions is to take carefully planned measures to prevent, improve, or treat psychological disorders or problems of a psychological nature (Ricou et al., 2019). An effective intervention aims to enhance individual resources, enabling the person to develop autonomy and adaptability in various situations. This involves equipping individuals with self-help skills to effectively manage and cope with the reoccurrence of symptoms related to their condition (Ricou et al., 2019). An effective intervention is grounded in evidence-based practices. Extensive research has demonstrated the efficacy of various psychological interventions for specific diagnoses. For instance, evidence-based interventions have been established for condi- tions such as depression (Hazell, 2021), addiction (Clay, 2008), anxiety (Turner & McCarthy, 2017), and schizophrenia (Bueno-Antequera & Munguía-Izquierdo, 2020). Evidence-based interventions are developed through a collaborative process that involves the exchange of information and close collaboration between researchers and practitioners.

###### Evaluation

Evaluation, which involves assessing the effectiveness of interventions, is a crucial process in various fields, including clinical, educational, and business sectors. In applied psychol- ogy, evaluation is conducted to examine the practicality of measures, identify areas for improvement, and make decisions regarding their implementation. The specific methods used in an evaluation depend on the application of research methods.

#### Hypotheses

Hypothesis formulation is a crucial component of scientific and research processes, partic- ularly in quantitative research methods. A hypothesis can be defined as a concise and pre- cise statement that is both testable and falsifiable, representing the researcher’s predic-

tion of the expected outcome of the study (Cuttler et al., 2019; McLeod, 2018). A hypothesis is a preliminary answer to a research question that is based on the existing state of research. It is formulated in a way that allows for empirical testing, meaning that abstract concepts (such as intelligence or emotions) can be operationalized and measured (Bhandari, 2022). Furthermore, a hypothesis is refutable and extends beyond the individ- ual, aiming for general validity. In scientific research, hypotheses are always grounded in existing theories and prior research. This implies that a scientific hypothesis cannot be for- mulated in the absence of previous scientific knowledge on a particular question, and assumptions cannot be derived solely from findings in related areas (Bhandari, 2022). In this scenario, the research question is approached in an exploratory manner. The four properties of theoretical justifiability, empirical testability, refutability, and general validity distinguish a scientific from a non-scientific hypothesis (Bhandari, 2022).

In a research study, it is common to have multiple hypotheses depending on the nature of the project and the research questions being investigated. These hypotheses often involve propositions about relationships between two or more variables, specifically the inde- pendent variable (IV; the variable manipulated by the researcher) and the dependent vari- able (DV; the variable measured by the researcher). By formulating multiple hypotheses, researchers can explore various potential relationships and outcomes within their study (Cuttler et al., 2019; McLeod, 2018). In any research study, it is essential to present the hypothesis in two forms: the null hypothesis (H0) and the experimental or alternative hypothesis (H1). The null hypothesis (H0) states that there is no relationship or difference between the variables under investigation, suggesting no effect. On the other hand, the alternative hypothesis (H1) asserts that there is a link between the variables, with one vari- able influencing the other. The alternative hypothesis specifies the expected changes in the DV as a result of the manipulation of the IV. These changes can be examined through statistical markers such as mean values, their differences, or associations, which are refer- red to as effect sizes. In statistics, a statistical association is known as a correlation. The alternative hypothesis also indicates that the findings are not due to chance and provide support for the theory being examined (McLeod, 2018). Additionally, the alternative hypothesis proclaims that the findings are not due to chance and are statistically signifi- cant, providing support for the theory being examined (McLeod, 2018). The null hypothe- sis (H0) , in contrast, asserts that there is no relationship or difference between the varia- bles being investigated, indicating no effect. It suggests that the DV, which is what the researcher is measuring, will not change as a result of the experimental manipulation imposed by the IV (McLeod, 2018). With respect to the findings, the null hypothesis states that the outcomes of the study were due to chance and do not provide support for the theory or concept being tested (McLeod, 2018).

An additional distinction to consider in relation to the hypothesis is its directionality. A hypothesis can be constructed as either directional or non-directional, but it cannot be both (McLeod, 2018). A directional hypothesis, also known as one-tailed hypothesis, pre- dicts the specific direction of the effect the IV has on the DV (greater or smaller). A non- directional hypothesis, also known as a two-tailed hypothesis, indicates that the IV influ- ences the DV without specifying the direction of the effect. It claims there will be a difference or an effect without suggesting a greater or smaller effect (McLeod, 2018).

An example of a null hypothesis is: “There is no relationship between the hours children watch TV and the violent incidents the same children show during play”. The alternative hypothesis would be: “There is a relationship between the hours children watch TV and the violent incidents the same children show during play”. According to Bandura’s Social Learning Theory (McLeod, 2016), children learn by observing and imitating the behaviors, attitudes, and emotional reactions of others. Based on this theoretical model and existing empirical evidence, it is known that early exposure to television can have negative effects on children’s development, particularly in relation to violence (Huesmann et al., 2003). Therefore, we will assume that the alternative hypothesis is correct, which we can also refer to as the working hypothesis. Regardless of whether our working hypothesis is the null or the alternative hypothesis, it is always evaluated using statistical testing, as this is essential for the application of statistical methods.

It is important to note that our hypothesis does not have a specific direction. We simply state that there will be a relationship between the number of hours children watch TV (one variable) and the occurrence of violent incidents during play (second variable). We do not specify whether watching more TV leads to more violent incidents or vice versa. Therefore, we would characterize our hypothesis as two-tailed or non-directional. However, the same hypothesis can be formulated differently if we wanted to have a one-tailed or directional hypothesis. In that case, we would write it as follows: “The more hours children watch TV, the more violent incidents children will show during play.”

The following table provides some examples of hypotheses to help illustrate the difference between scientific and non-scientific hypotheses. The concept of DVs and IVs will be cov- ered in more detail in the next subsection on variables.

Table 1: Differences between Non-Scientific and Scientific Hypotheses

|  |  |  |
| --- | --- | --- |
|  | **Non-scientific hypothesis** | **Scientific hypothesis** |
| Properties | from gut feeling, traditions etc. | from the current state of research (the- ories, empirical findings) |
| Example | Zodiac signs influence personality traits. | Knowledge about the zodiac signs influences self-assessed personality traits. |
| Explanation | So far, we have no empirical findings that the date of birth or zodiac sign influences personality traits. | People who are interested in zodiac signs and know typical traits of the zodiac sign rate themselves according to these traits. |
| Properties | not testable | testable |

Source: Evangelos Zois (2022).

#### Variables

The main characteristic of hypotheses is that they are testable, which refers to the ability to operationalize and measure all the relevant characteristics included in the hypothesis. In psychology, some characteristics can be directly measured, such as reaction time in mil- liseconds, age in years, or points on a school test. However, researchers often aim to measure abstract phenomena that cannot be directly observed, such as personality, iden- tity, or happiness. To capture these abstract elements, various techniques are used in psy- chological research, including self-report questionnaires, behavioral observations, psy- chological tests, and tasks. The quantities collected using these techniques are referred to as variables, which play a crucial role in psychological research. Variables can be manipu- lated or take on different characteristics or values, allowing researchers to examine their relationships and effects (Goodwin, 2010; Howell, 2010).

##### Manifest and Latent Variables

When a variable can be directly measured or counted, it is referred to as a manifest varia- ble. Manifest variables are assumed to have no measurement error. Examples of manifest variables include reaction time, sleep duration, and salivary cortisol level. On the other hand, abstract phenomena in psychology that are measured indirectly are called latent variables. For instance, family functioning level is a latent variable that can be measured using ten questions answered on a three-point scale. Each of these individual questions represents a manifest variable. These manifest variables are used to capture the underly- ing latent variable of family functioning level. Since latent variables often represent com- plex constructs, it is assumed that there is some degree of measurement error. Efforts are made to minimize this measurement error in order to obtain reliable characteristics. The distinction between latent and manifest variables is particularly relevant in certain quanti- tative research methods, such as factor analysis (Hand, 2004).

##### Independent and Dependent Variables

As mentioned earlier, the categorization of variables into independent variables (IV) and dependent variables (DV) is crucial in the design of a quantitative research project as it determines what will be measured, how it will be measured, and how changes in the vari- ables may occur during an experiment (Kranzler, 2018). An IV is a variable that the researcher or experimenter manipulates in order to examine its effect on another variable (Kranzler, 2018). The DV refers to the measure being taken by the researcher or experi- menter. The expected change in the DV is due to the manipulation of the IV in the experi- ment. Consequently, the DV is the primary focus of any statistical analysis conducted (Kranzler, 2018). For example, imagine a hypothetical research scenario in which a researcher examines whether gender affects the expression of emotion regulation. The goal is to measure emotion regulation and then assess whether its expression differs when introducing gender as a variable. The DV, or what is measured, is emotion regulation, and the IV, or what is manipulated, is gender. Gender will remain constant in response to var- ied approaches to emotion regulation, but the approach to emotion regulation is likely to be influenced by gender.

##### Continuous and Discrete Variables

Another important distinction in regards to the characteristics or values variables can take is that of continuous and discrete variables. Discrete variables can only take on specific values, such as the number of individuals in a family, which can be 2, 3, 4, 5, and so on. It is impossible to have 2.4 people in a family. Continuous variables, however, can take on any value such as height, weight, and age. It is possible to be 72.4 kilograms or 24.5 years old (Kranzler, 2018).

##### Scales of Measurement

Measurement involves assigning numbers to the characteristics or variables on which indi- viduals vary. These numbers reflect the variability in people’s characteristics, allowing for a quantitative representation of those variations. However, different types of variables have different rules for assigning numbers that accurately indicate how individuals differ on those variables (Kranzler, 2018). There are four different scales of measurement that variables can be measured on: nominal, ordinal, interval, and ratio. Each scale has specific rules for assigning numbers to variables based on their properties. These properties of variables at different scales are crucial in determining which methods to use when analyz- ing them (Kranzler, 2018).

###### Nominal Scale of Measurement

Variables that belong to the nominal scale of measurement are often referred to as catego- rical variables. This scale involves assigning individuals or objects to specific categories that explain how they differ, but without any numerical meaning. Examples include gen- der (male, female), eye color (brown, blue, green), race (black, white), or sexuality (hetero- sexual, homosexual, bisexual). People who belong to the same category are considered equal, but the numbers used to label the categories do not imply any order, size, or value. The numbers assigned to each category are arbitrary and serve as labels rather than indi- cating any inherent numerical significance. For instance, using 1 to represent male and 2 to represent female is simply a way of categorizing individuals and does not imply any numerical relationship or hierarchy between the categories (Kranzler, 2018).

###### Ordinal Scale of Measurement

For this type of measurement, magnitude is key. Objects or people are ranked on a contin- uum based on their magnitude. Numbers are then used to create an ordered ranking of individuals or objects (Kranzler, 2018). Examples of variables used in this type of ranking include socioeconomic status (low income, middle income, high income), education level (high school, bachelor’s degree, master’s degree, doctorate), and marathon rankings (1st place, 2nd place, 3rd place). For the education level variable, a rank is assigned with High School graduates given the number 1, BSc graduates given the number 2, and MSc and PhD graduates given numbers 3 and 4 respectively. This rank expresses a “greater than” association (Kranzler, 2018). However, it does not provide information on the extent of the difference. While we understand that a BSc degree signifies higher knowledge compared to a high school degree, the exact degree of difference in terms of knowledge cannot be determined. To illustrate this, consider the example of a marathon: the runner who fin-

ishes first is assigned number 1 and the runner who finishes second is assigned number 2. While we can ascertain who finished first, we do not know the magnitude of the difference between the first and second runners. The difference could be as small as 5 seconds or as large as 5 minutes (Kranzler, 2018). We cannot assume that the difference between the first and second ranks is equal to the difference between the second and third ranks, and so on. This lack of uniformity in comparative merit is not reflected in the assigned ranks.

###### Interval scale of measurement

The numbers assigned to variables in this scale of measurement indicate a comparative merit and have equal units of measurement. This means that the same difference between two points on the scale holds the same significance in terms of the measurement being considered (Kranzler, 2018). Examples of variables belonging to the interval scale of measurement include age, temperature, SAT scores, and IQ test results. This scale of measurement lacks an absolute zero, meaning we cannot calculate ratios between scores. For instance, let us consider the IQ test. If one person scores 100 on an IQ test and another scores 90 on the same test, we can objectively state that person A is more intelligent than person B. If we examine the score of a third person who scored 80 and a fourth person with a score of 70, we can logically conclude that the 10-point difference between person C and person D is the same as the 10-point difference between person A and person B. It is important to note that with an interval scale, a true zero does not exist; therefore, we can- not have zero intelligence or zero temperature, as their absence does not equate to zero. A final note about ratios: an individual with an IQ score of 100 is not twice as intelligent as someone with a score of 50, nor is a temperature of 20 degrees twice as warm as 10 degrees (Kranzler, 2018).

###### Ratio scale of measurement

Variables measured on a ratio scale of measurement possess a true zero value, indicating an absolute absence or nothingness, and equal units of measurement (Kranzler, 2018). Examples of ratio scale variables include height, weight, reaction time, and calories. As the name suggests, the ratio scale allows for the calculation of ratios between values. For instance, we can confidently state that 10 kilograms is half as heavy as 20 kilograms, or that an individual weighing 100 kilograms is twice as heavy as another individual weighing 50 kilograms.

Table 2: Determining the Scale of Measurement of a Variable

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Scale** | **Category** | **Order makes sense in terms of content** | **Calculating makes sense in terms of con- tent: the same distance between the values** | **Absolute zero** |
| Nominal | **✓** | **x** | **x** | **x** |
| Ordinal | **✓** | **✓** | **x** | **x** |
| Interval | **✓** | **✓** | **✓** | **x** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Scale** | **Category** | **Order makes sense in terms of content** | **Calculating makes sense in terms of con- tent: the same distance between the values** | **Absolute zero** |
| Ratio | **✓** | **✓** | **✓** | **✓** |

Source: Evangelos Zois (2022).



**SUMMARY**

Psychology is an empirical discipline wherein knowledge is acquired through observation and analysis of data. The objective is to achieve comprehensive, systematic, and clear knowledge. The methods used for the collection and evaluation of data in psychology can be quantitative, qualitative, or mixed. In quantitative research, data is gathered and assessed through standardized psychological evaluations, experimental tasks, questionnaires, and interviews. Participant responses are stored and coded in a consistent manner. Qualitative research, on the other hand, involves gathering and analyzing unstandardized data. Mixed methods research integrates both quantitative and qualitative strat- egies. Quantitative research follows a deductive approach that tests sci- entific hypotheses using empirical data. In contrast, qualitative research collects and analyzes empirical data to derive new theories and hypoth- eses through an inductive approach. Formulating hypotheses and oper- ationalizing variables are essential elements in any research process. Scientific hypotheses are derived from the current state of research, and they are testable, refutable, and generally valid. Hypotheses can be either directional or non-directional depending on the research ques- tion. The variables chosen to build the hypotheses have varying levels of measurement: nominal, ordinal, interval, and ratio. However, it is most crucial to differentiate between independent and dependent variables, which represent what the experimenter manipulates and what the experimenter measures, respectively.

# UNIT 2

## PROBLEM AND RESEARCH CYCLE

**STUDY GOALS**

On completion of this unit, you will be able to ...

* explain how research projects in psychology are developed and implemented.
* describe how scientific questions emerge.
* provide a comprehensive overview of various sampling methods employed in research.
* describe the factors that determine the representativeness of a sample.
* describe available data collection techniques.

### 2. PROBLEM AND RESEARCH CYCLE

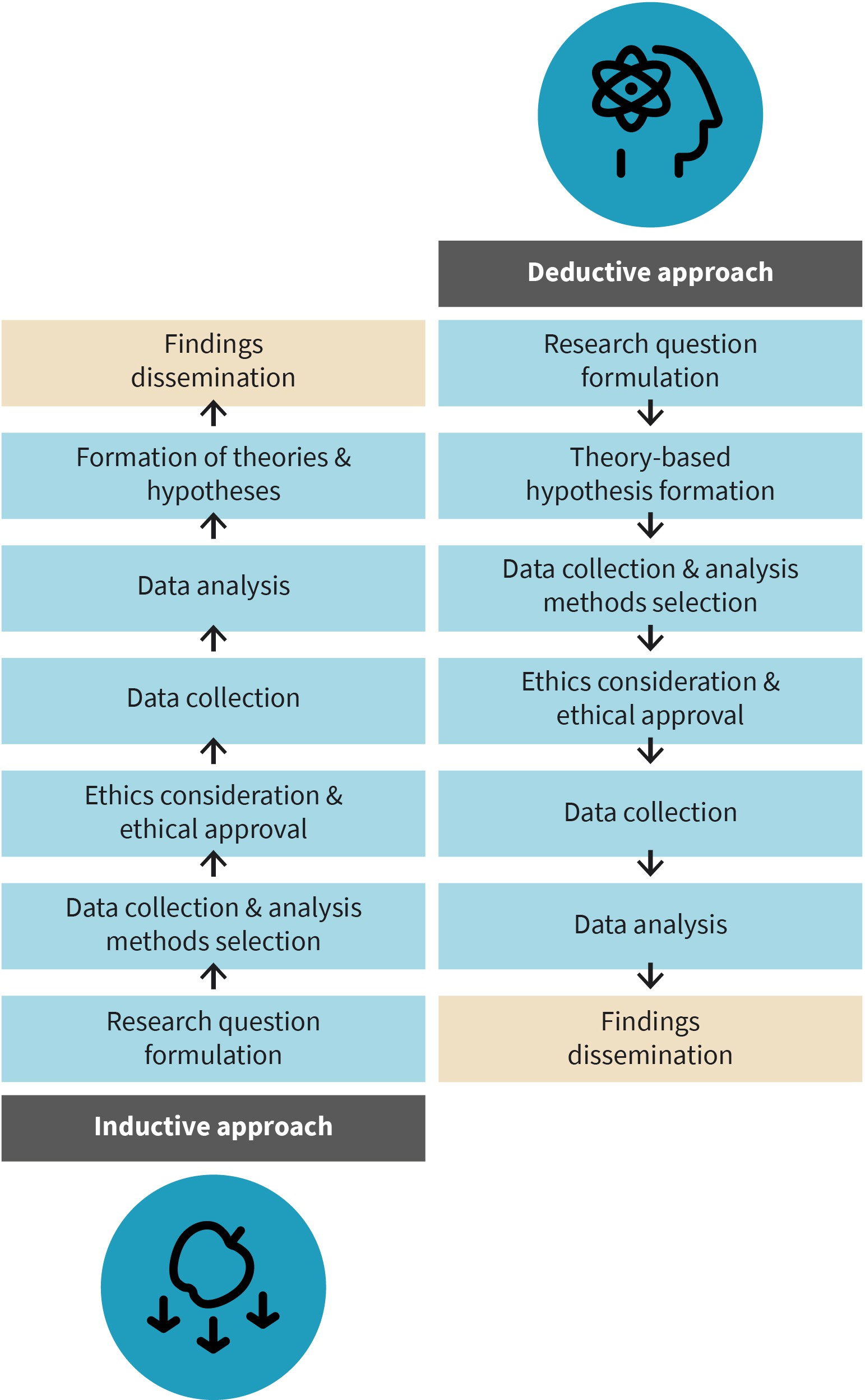
#### Introduction

As a psychology student at the IU, you will be required to select a research thesis topic and conduct an independent empirical project. However, you may be wondering how to approach this process from a scientific standpoint. How do you choose a research subject and develop your own research questions? What steps should you take before collecting and analyzing data? This unit will guide you through the research process involved in psy- chology research projects. Additionally, you will learn about the primary techniques used in sampling and data collection.

#### Research Cycle

Scientific studies and projects adhere to a well-defined sequence of steps, beginning with the conceptualization of the research idea and concluding with the collection and analysis of data. Undergraduate projects or theses, while smaller in scale, also adhere to these guidelines and must follow specific steps. It is important to note that each research endeavor may adopt a different approach, either deductive or inductive, and employ either qualitative or quantitative research methods. However, the overall process remains consistent. The graphic below illustrates the typical flow of the research process, featuring separate illustrations of the deductive and inductive approaches.

Figure 3: A Schematic Representation of the Research Process



Source: Evangelos Zois (2022).

Ethics play a crucial role in research involving human participants, and it is essential to give them serious consideration. Ethical considerations should be thoroughly examined at every stage of the research process, including planning, data collection and analysis, and the dissemination of findings. Best practices in scientific research recommend that ethical aspects of a project should be addressed prior to the initiation of data collection. For any study involving human subjects, it is imperative to obtain approval from an official ethics committee. Such committees are established in academic and research institutions to evaluate proposed research projects and determine their justifiability and potential risks to participants. The committee closely scrutinizes every aspect of the research, including the selection of research questions, available resources and personnel, and the proposed methods for data collection and analysis. The primary focus of every research inquiry is the safety and well-being of participants who voluntarily take part.

##### Selection of a Research Topic

The first step in initiating a research project is to carefully select and decide on a specific topic to investigate, followed by the formulation of research questions. Choosing a research topic is not a simple task and even experienced researchers may find it challeng- ing. It is crucial to consider personal interest when selecting a topic. Additionally, it is important to have a good understanding of the chosen topic and be aware of the existing body of research on that topic. This knowledge allows an identification of gaps in the field that researchers can contribute to by providing new insights to the scientific community. Following this, a set of relevant principles are applied to guide the process of selecting a research topic (Howitt & Cramer, 2017; Cuttler et al., 2019).

###### Personal interest

Embarking on a research project will inevitably reveal that the research process can be demanding and time-consuming. Therefore, it is advisable for researchers at any level to choose an interesting topic to investigate. By doing so, they can maintain their motivation throughout the arduous research journey. In the field of psychology, it is common for researchers to explore topics that hold personal significance. For instance, if a researcher has personal experience growing up in an environment with addiction issues, they may be inclined to empirically study addiction in order to gain deeper insights into the condition.

###### Existing theoretical and methodological knowledge

The topic of investigation should be relevant to the researcher’s existing knowledge and qualifications. For example, psychology students typically choose a topic closely aligned with their field of study or related disciplines. Nowadays, research projects often involve collaboration and interdisciplinary efforts. This means that various scientific experts, such as statisticians, psychologists with clinical training, and physicians, may be involved depending on the complexity and scale of the study. It is necessary for researchers to be well-informed and aware of the current state of research in order to identify gaps and jus- tify the need for further investigation in the area.

Consider the case of a young researcher interested in studying anxiety disorders and gen- der differences. Initially, they propose an examination of whether there is a gender differ- ence in anxiety, but their ethical approval is rejected due to existing published research indicating no gender difference in anxiety. Upon further reading, the researcher realizes that previous studies have not explored the potential role of age in relation to anxiety. In response, they reformulate their research question and resubmit it for approval, since a gap in the literature has been discovered.

Another crucial aspect of good scientific practice is the extent to which a proposed research study contributes to the field. This does not necessarily involve exploring entirely new topics, but rather researchers can initiate studies to replicate existing findings. Vali- dating previous important findings is highly significant in modern psychological science. When we replicate previous findings, we enhance the validity of those findings. Replica- tion implies delivering the same or similar results as the original investigation. This increases the confidence in generalizing our results to the broader population. In recent years, psychology has faced a replication crisis, as researchers struggle to reproduce con- sistent results from previous studies (Maxwell et al., 2015). The reasons for this crisis are beyond the scope of this section, hence it will not be discussed further. If you are inter- ested in learning more about this topic, please refer to the suggested reading list provided.

When selecting a research topic, it is important for researchers to consider its practical relevance and the potential implications of the findings. For instance, in the field of psy- chology, if a researcher investigates workplace anxiety and finds that there is an increase in anxiety with age, regardless of gender, this information can be valuable for applied psy- chology. An applied research psychologist can explore ways to reduce workplace anxiety specifically for older employees. On the other hand, basic research focuses on under- standing psychological mechanisms without a direct application of findings. However, basic research findings indirectly contribute to practical and applied psychological research, as demonstrated in the example of anxiety in the workplace.

Another significant factor to consider is the capability of investigating the chosen topic. Researchers may encounter limitations that prevent them from examining certain topics. These limitations can relate to

* ethical concerns raised by the study (e.g., if violence hinders memory recall),
* inaccessibility of the target group,
* the need for a large number of participants resulting in high recruitment costs, or
* the unavailability of necessary data collection tools.

For instance, if a researcher is interested in studying how individuals who have committed violent acts continue to lead seemingly normal lives without being apprehended, they will face significant challenges. It is nearly impossible to locate and approach such individuals as they would be inaccessible and unlikely to willingly participate in the study.

The importance of conducting extensive and thorough research before formulating a research question has already been emphasized. Research involves reading and studying specialized books and published journal articles. To find relevant published work, litera- ture searches should be conducted using various options. Specialist research articles can

be found in specialized journals. A comprehensive list of psychology-related journals can be online. Instead of searching each journal individually, electronic journal databases such as “Web of Science,” “PsycInfo,” “Psyndex,” and “PubMed” offer a more efficient way to search for desired information.

To facilitate the rapid dissemination of research findings, many manuscripts of research studies are now available online as preprints. These preprints are made available before the articles undergo the traditional process of peer review and formal publication. As an example, in February 2020, just two months after the initial identification and rapid spread of Covid-19, one of the first articles examining psychological stress and the use of social media during quarantine was made available as a preprint (Gao et al., 2020). Findings from preprints should be interpreted cautiously since they have not yet undergone the process of external feedback and peer review. Peer review is a crucial step in research pub- lication, where manuscripts are evaluated by experts in the field prior to official publica- tion. Once a researcher has completed a study and analyzed the data, the next step is to disseminate the findings by publishing them in a scientific journal. Based on the topic of investigation, the researcher selects a journal, such as those listed previously or other psy- chology-related journals, and submits the manuscript for publication. The manuscript is formatted as a research report and undergoes the peer review process. This entails send- ing the manuscript to two or three experts in the field, who evaluate it for potential errors, seek clarifications, and ensure that the study adheres to ethical guidelines. These experts ultimately determine the outcome of the manuscript, deciding whether it should be pub- lished or not.

##### Method Selection and Application

After identifying the research objective, the researcher must determine the appropriate method to employ, which could be either inductive or deductive. For example, if a theory suggests that fear-related information contributes to the development of phobias and this theory has been confirmed by studies conducted in the adult population, the deductive approach would be suitable. On the other hand, the inductive approach would be appro- priate if no existing theory on the topic is available. The next step involves selecting the appropriate methods for data collection and analysis, which depends on the nature of the research investigation. The researcher should have a broad understanding of the advan- tages and disadvantages of different methods, whether qualitative or quantitative. Follow- ing this, the researcher will proceed with the preparation of necessary materials and docu- ments for the study, such as information sheets, consent forms, and data collection tools. The data is then collected, and it is crucial to meticulously document the process in order to identify and consider any potential errors when analyzing the data. Once data collec- tion is completed, it should be saved and stored in a suitable format. For quantitative data analysis, software such as Excel, IBM SPSS, or R can be utilized. In the case of qualitative methods, the data is transcribed and standardized to facilitate analysis (Howitt & Cramer, 2017).

Quantitative data analysis involves the application of statistics and mathematical models. Statistics can be broadly classified into two main areas: descriptive statistics and inferen- tial statistics. Descriptive statistics focus on summarizing and describing the main charac- teristics of the study data, providing a better understanding of the data. However, descrip-

tive statistics alone cannot draw conclusions about the study or its findings. In contrast, inferential statistics are used to make inferences and draw conclusions about a population of interest based on the investigation of a representative sample. These statistics involve the use of various tools and statistical methods to analyze data, allowing researchers to make statistically valid inferences and generalize findings to the larger population (Kran- zler, 2018).

##### Interpretation of Findings

The interpretation of findings involves several steps. Firstly, the researcher compares the results obtained to the initial hypotheses proposed. Based on statistical analysis, a deci- sion is made regarding which hypotheses to accept or reject. Next, the findings are exam- ined in relation to previous research conducted in the same field. If the original hypothesis is not confirmed and must be rejected, the researcher discusses possible reasons for this and considers potential limitations of the study that could have influenced the results. Conversely, if the results align with the initial hypothesis, the researcher discusses the results in relation to previous findings. Lastly, the researcher explores the implications of the findings for future research in the field and discusses any practical applications or implications that may arise from the results (Howitt & Cramer, 2017).

##### Presentation of Results

The final step in the research process involves summarizing the results for dissemination through the preparation of a scientific report, such as a journal article or a degree thesis. The scientific report should typically include the following sections presented in the described structure. The same structure should be followed when preparing a degree the- sis (Howitt & Cramer, 2017).

###### Abstract

Although this is the first section of a scientific report, it should be written at the end. It follows a standard format, providing a brief overview of the background and objectives of the research investigation, outlining the methodology used (including sample, data col- lection methods, and analysis techniques), presenting the main findings, and concluding with their implications. Abstracts are concise, typically ranging between 150 to 250 words.

###### Introduction

This section presents an introductory overview of the topic, including the relevant theo- retical background and state of research. The aims of this investigation are described in detail, with an explanation of why this topic deserves examination and what gaps in knowledge remain. Following this, the research question(s) are clearly stated. The section concludes with a concise declaration of the hypotheses, as needed for the study. It is important to keep in mind that each hypothesis has a corresponding null and alternative/ experimental hypothesis.

###### Methodology

This section presents detailed information about sample (size, characteristics, sampling method, etc.), data collection procedure, data collection tools (questionnaires, tests, etc.), and statistical analysis methods (software, statistical p-values, etc.).

###### Results

This section provides a thorough and systematic presentation of the findings, encompass- ing both descriptive and inferential analyses. Findings are often accompanied by graphical representations and tables. It is important to note that the focus of this section is solely on presenting the findings themselves, without providing any interpretation or explanation. The interpretation and discussion of the findings are reserved for the subsequent section.

###### Discussion

The final section of the scientific report should be of similar length to the introduction. the researcher begins by discussing which hypotheses are accepted or rejected based on the findings. The results are then contextualized within the current body of research in the field. The researcher examines whether the findings align with previous studies and, if not, provides possible explanations or reasons for divergence. The following paragraph is dedi- cated to acknowledging and addressing the limitations of the current research (e.g., draw- backs that should be considered for future research on the same topic). Suggestions for future research directions may be proposed in the same paragraph. Finally, a general con- clusion is presented, focusing on the practical and scientific implications of the research findings.

**Representative sample** The representative sam- ple refers to the group under investigation that closely resembles the population of interest (in terms of location, age, income, gender, etc.). The sample should mirror the population that is drawn

from.

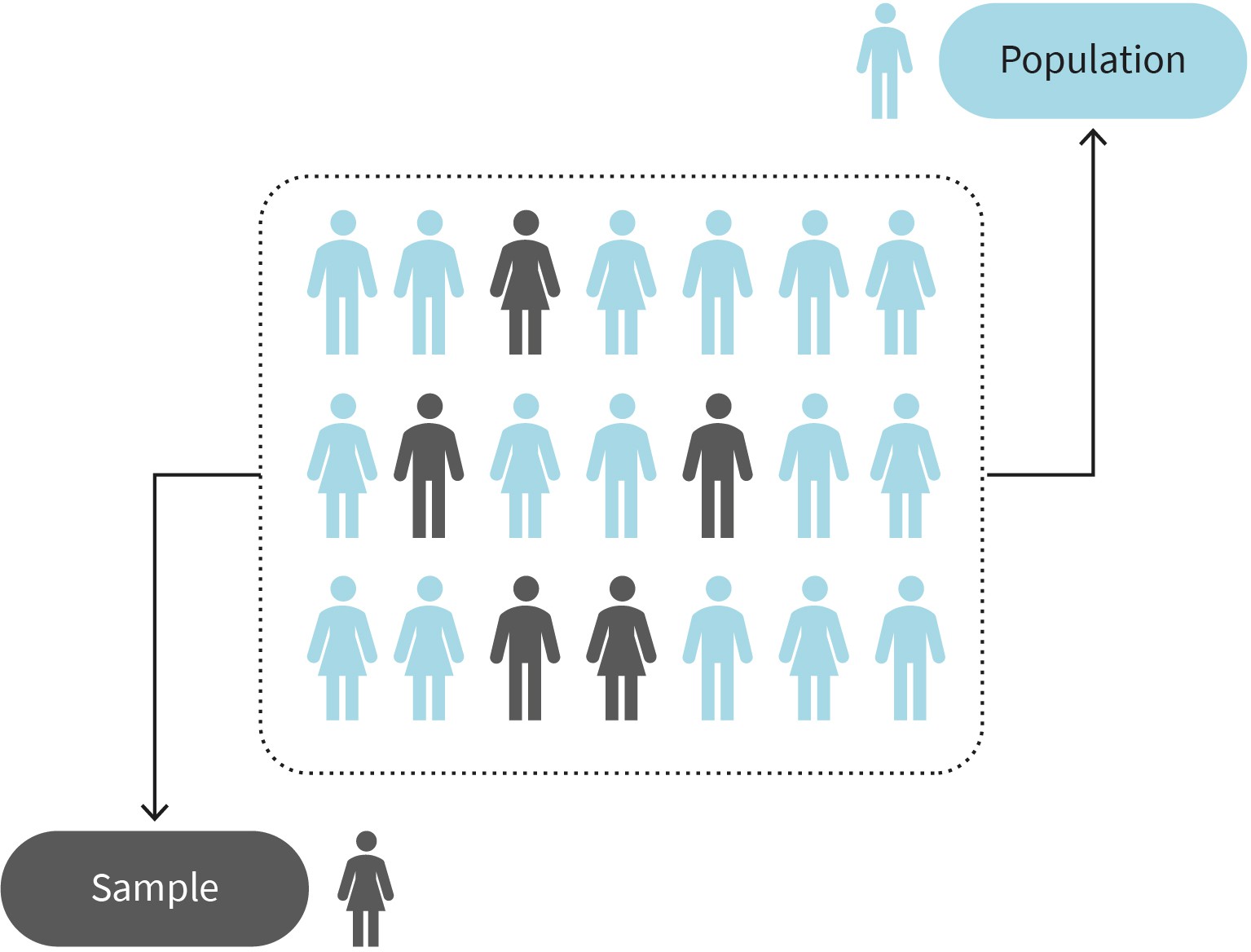
#### Sampling

This section discusses the concept of sampling provides an overview of the most com- monly used sampling methods in psychological research. To illustrate the importance of sampling, consider a scenario where a large pharmaceutical company aims to investigate the adverse effects associated with a new medicine. Since the company is based in the EU, their focus would be on testing the effects of the medicine on the European population. However, it would be impractical to include every EU citizen in the research investigation. Therefore, the company would need to employ a suitable sampling method to ensure they obtain a **representative sample** of the EU population for their research project. Careful selection of the sample is crucial to ensure its representativeness of the larger EU popula- tion.

A sample refers to a specific group of individuals who participate in a research project (McCombes, 2019). When conducting research that requires sampling, it is crucial for the researchers to carefully select a sample that is representative of the population of interest. This is essential to ensure that valid conclusions can be drawn from the results obtained. The population, on the other hand, encompasses the entire group of people that the researchers aim to examine, such as the EU in the case of the pharmaceutical company.

It is important to note that the sample size is always smaller than the total population of interest (Howitt & Cramer, 2017; Cuttler et al., 2019; McCombes, 2019). To provide further clarification, some simple examples can be used to differentiate between population and sample. For the pharmaceutical company, all EU countries would constitute the popula- tion of interest, while the sample would consist of the countries where the company has branches, such as Germany, Switzerland, and Spain. In another example, the population of interest would be employees of Microsof t Germany, while the sample would be limited to Microsof t employees in the city of Berlin only. Lastly, the target population may be undergraduate psychology students in Germany, with the sample consisting of 500 stu- dents selected from five German universities.

Figure 4: Population and Sample



Source: Evangelos Zois (2022).

Sampling is a crucial aspect of research because it allows us to study a subset of the popu- lation instead of attempting to include and analyze the entire population. The population size is often too large to feasibly recruit and manage, and even if it were possible, not all individuals in the population would be accessible for various reasons. Therefore, examin- ing the entire population is impractical, unmanageable, and not cost-effective (Jhangiani et al., 2019).

##### Sample size

The sample size depends on the population size. If the target population is large, a larger sample size is required (Foster et al., 2018; Martínez-Mesa et al., 2014). In contrast, a small target population will require a smaller sample size. However, it is always essential to ensure that the sample size is adequate to represent the target population accurately. A minimum sample of 50 participants is adequate for most psychological experiments (Brys- baert, 2019). However, for surveys, a larger sample, typically around 150 participants, is recommended (Marszalek et al., 2011). Another approach suggests that the sample size should be approximately 10% of the population size, as long as it does not exceed 1,000 participants (Bisits, n.d.). There are sophisticated software packages available that can assist researchers in determining the optimal sample size based on various factors. How- ever, these tools are often more relevant and utilized in advanced stages of research rather than in undergraduate-level research.

##### Types of Sampling Methods

**Random selection** Every individual within the population of interest has an equal opportunity

to be selected.

The sampling methods which are commonly used in psychological research are divided into two types: probability sampling and non-probability sampling (Foster et al., 2018; Howitt & Cramer, 2017; Cuttler et al., 2019). This learning unit will center on eight distinct sampling methods. The initial four methods will be probability sampling techniques, while the remaining four methods will fall under non-probability sampling approaches. The key distinction lies in the fact that probability sampling methods involve **random selection**, enabling robust statistical inferences to be made (McCombes, 2019). Non-prob- ability sampling is a simpler approach for data collection that does not involve random selection (Howitt & Cramer, 2017).

###### Probability sampling methods

The following belong to probability sampling methods:

* **Random sampling** gives each element within the population an equal chance of being included in the sample. This type of sampling can be implemented through computer- based tools that generate random numbers. For example, when examining satisfaction among members of a gym with a total membership of 1,000 and a desired sample size of 100, each member can be assigned a number from 1 to 1,000. Subsequently, a ran- dom number generator can be utilized to select 100 random numbers from the popula- tion. Random sampling is advantageous due to its relatively simple implementation and the minimized potential for bias. Researchers often favor this method to generalize findings to the population of interest. However, there are drawbacks to consider. Ran- dom sampling can be time-consuming and expensive, particularly when dealing with larger populations. Despite researchers’ efforts to minimize bias, it may still be present in the data (Foster et al., 2018).
* **Cluster sampling** divides the population into subgroups, with each subgroup ideally possessing characteristics similar to the target population. Rather than sampling indi- viduals from each subgroup, the subgroups themselves are randomly selected (McCombes, 2019). For example, a school principal interested in gathering student opin- ions on the administration process may choose specific classes to obtain responses. All

students within the selected classes will have the opportunity to participate. Cluster sampling is a cost-effective method and a practical way to include more participants in a research study. However, a disadvantage of cluster sampling is that it can sometimes be challenging to obtain clusters that accurately represent the entire population, result- ing in sampling error (Foster et al., 2018).

* + **Systematic sampling** is akin to random sampling but is generally perceived as easier to implement. In this method, every element of the population of interest is assigned a number, similar to random sampling. However, unlike random sampling, systematic sampling does not involve randomly generated numbers. Instead, individuals are selected at regular intervals (McCombes, 2019). For example, consider a scenario where all gym members are listed alphabetically. The first 10 numbers are selected randomly, resulting in a gym member being assigned the number 5. Starting from this point, every 10th person on the alphabetically generated list is chosen. This process continues, selecting the 5th, 15th, 25th, 35th member, and so forth until the desired sample size is reached. Systematic sampling is a straightforward method to implement, and it reduces the likelihood of bias in the data. However, a potential disadvantage is the possibility of over- or underrepresentation of certain characteristics within the final sample (Foster et al., 2018).
  + **Stratified sampling** divides the population into subpopulations (strata) based on rele- vant characteristics such as gender, income, or age. Subsequently, the researcher deter- mines the sample size for each subgroup by considering the proportions present in the overall population (McCombes, 2019). Finally, the process of systematic or random sam- pling is employed to select the sample from each subgroup. For instance, consider a university psychology course with 800 female students and 200 male students. To cre- ate a sample that is representative of the gender discrepancy in the population, the researcher would separate the population into two strata based on gender. Assuming a desired sample size of 100 participants, random sampling would be used to select 80 female and 20 male student participants. Stratified sampling is advantageous as it includes subgroups (strata), enabling the generation of a sample that closely represents the population of interest. However, analyzing data obtained through this sampling method can be challenging (Foster et al., 2018).

###### Non-probability sampling methods

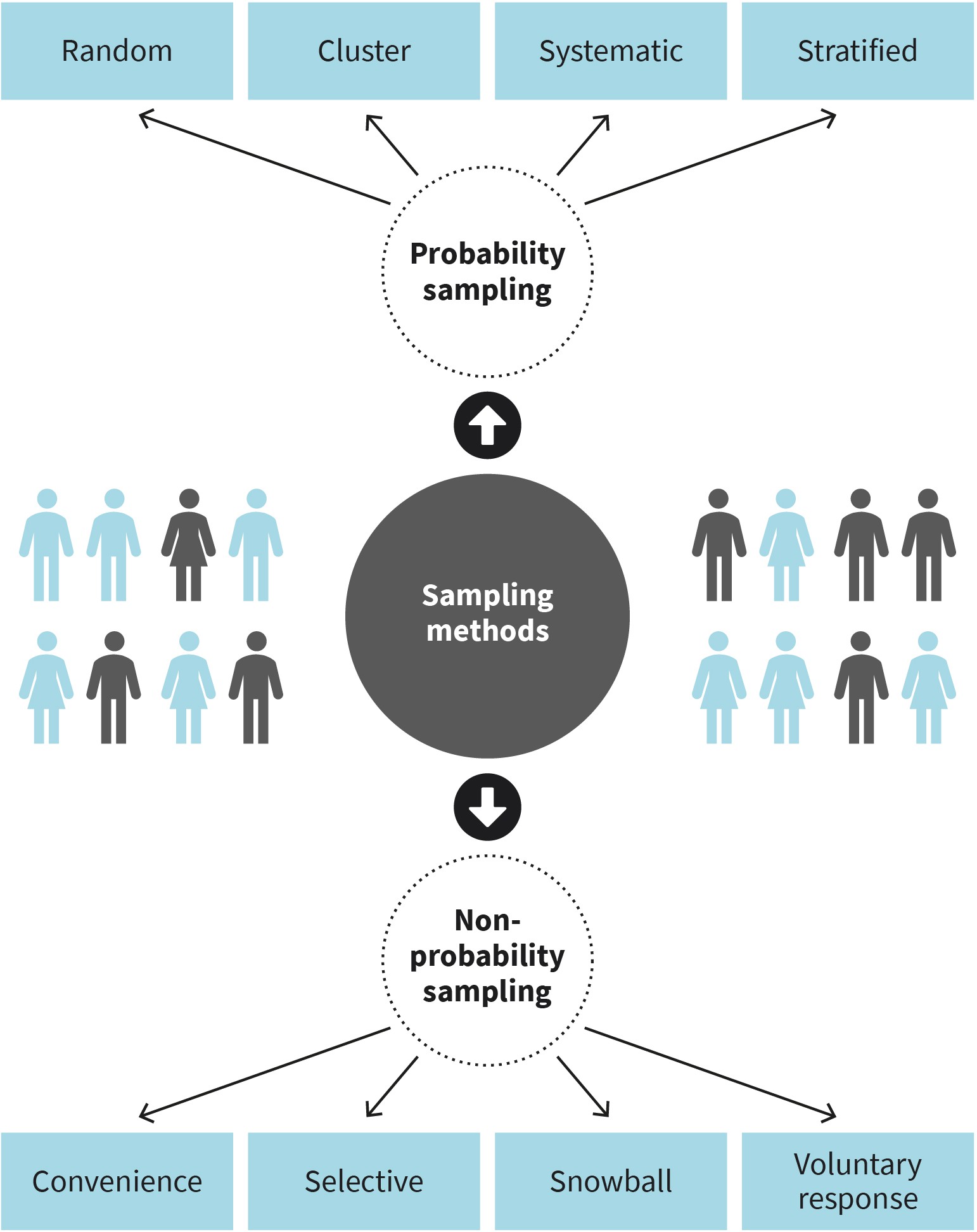
Non-probability sampling methods include:

* + **Convenience sampling** involves selecting people who are readily accessible to the research team during the recruitment process. This method is convenient, cost-effec- tive, and simple to implement. However, it is important to note that convenience sam- pling may not guarantee the sample’s representativeness, making it challenging to gen- eralize findings to the wider population (McCombes, 2019). For instance, a student conducting an undergraduate thesis may opt for convenience sampling by asking friends and family members to participate in a survey. While this approach offers advan- tages in terms of simplicity, efficiency, and affordability, its main drawback is the limited ability to generalize findings beyond the sample itself (Foster et al., 2018).
  + **Selective sampling**, also known as purposive or judgment sampling, is a type of non- probability sampling method where the researcher, using their expertise, selects a sam- ple that aligns with the specific needs and objectives of the research (McCombes, 2019).

This sampling technique is commonly used in qualitative research. It is important to note that in selective sampling, the chosen sample must have precise criteria and a jus- tified rationale for inclusion (McCombes, 2019). An illustrative example of selective sam- pling is when TV reporters on the street approach certain individuals for their opinions on a political situation. This method is cost-effective and inexpensive as every individ- ual in the population is considered suitable for participation. Nevertheless, selective sampling is susceptible to bias (Foster et al., 2018).

* + - **Snowball sampling** involves recruiting participants through the assistance of existing participants. This method is typically employed when the population of interest is chal- lenging to access (McCombes, 2019). For instance, consider a researcher interested in studying addiction rates among the homeless population in a city where authorities do not maintain a comprehensive list of homeless individuals. Fortunately, the researcher encounters a homeless person who expresses interest in participating and is willing to connect the researcher with other homeless individuals they know in the city who may also be interested in taking part. This method allows for a swift and cost-effective means of obtaining the required sample. Additionally, it increases the likelihood of including participants who would be hesitant to directly engage with a researcher. Some individuals may prefer to maintain their anonymity and avoid direct contact with the research team, even if they find the study personally relevant. However, one of the main disadvantages of snowball sampling is the potential for sampling bias (Foster et al., 2018).
    - **Voluntary response** is a non-probability sampling method which is similar to conven- ience sampling. The creation of the sample is based on easy access to participants. The researcher does not directly select participants but rather individuals volunteer by com- pleting an online survey that is available to everyone. However, this sampling method is not without bias since certain individuals may be more inclined to participate for vari- ous reasons. If an online survey is designed to examine people’s views on vaccination, it is possible that individuals with strong opinions against this medical practice will respond. While valuable information may emerge, researchers cannot be certain that the opinions recorded are representative of the population. The method is cost-effec- tive and simple to implement, but the researcher cannot control the sample’s character- istics (Foster et al., 2018).

Figure 5: Sampling Methods



Source: Evangelos Zois (2022).

###### Sampling bias

Probability sampling was primarily developed to prevent sampling bias. Sampling bias occurs when a sample is constructed that does not accurately represent the target popula- tion and delivers inaccurate results (Foster et al., 2018; Howitt & Cramer, 2017; Cuttler et al., 2019). When selecting a sample, researchers should try to avoid sample biases, which may include:

* + - **Undercoverage bias** occurs when not all individuals from the population are included in the sample, rendering it unrepresentative. For instance, when individuals are invited to participate in an online survey, those without internet access will not be included, leading to an unrepresentative sample.
    - **Non-response bias**, which means that it is common for individuals invited to partici- pate in a survey to not respond for various reasons, such as being deceased, busy, unin- terested, or generally not inclined to participate in surveys. When non-responders sys- tematically differ from the responders, it can lead to non-response bias (Cuttler et al., 2019). For example, in a study on alcohol consumption following treatment for alcohol use disorder, individuals who completed their treatment were contacted six months later to assess their drinking habits. However, only half of them responded, potentially introducing non-response bias. The individuals who did not respond may have different drinking patterns compared to those who did, which can significantly impact the study results. Researchers must actively seek ways to minimize non-response bias and improve response rates (Cuttler et al., 2019).
    - **Selection bias** arises when the selected sample does not accurately represent the pop- ulation of interest (target population). This is particularly common when the sampling method utilized is not based on random sampling, such as in the case of non-probabil- ity sampling methods like convenience sampling. Furthermore, high dropout rates (attrition bias) or non-responses may occur due to various factors associated with pro- longed participation in a research study. These instances should not be disregarded, but rather investigated to understand the reasons behind the dropout or non-response, as they may indicate that the selected population is not suitable for the study. Volunteer bias is another important consideration, as individuals who volunteer to participate in a study may possess traits that differentiate them from those who choose not to partici- pate initially, potentially rendering the responders unrepresentative of the population (Cuttler et al., 2019).
    - **Missing responses** means that during an investigation, individuals who initially agreed to participate may drop out before completing the study, leading to missing information in the dataset. In such cases, the researcher must conduct a thorough examination of the dataset to identify any errors that could potentially impact the study’s findings and their generalizability.

Researchers aim to recruit a representative sample that is not compromised by sampling bias whenever feasible. When designing a research study, it is crucial to have a clear understanding of how and who will be recruited, and to consider strategies that can enhance and maintain participant motivation. To achieve this, it is important to develop an appealing study title and a straightforward yet comprehensive description of the study.

#### Data Collection Techniques

In psychological research, the choice between qualitative, quantitative, or mixed methods approaches depends on the research topic and the specific research needs. Similarly, when referring to data collection techniques, it is not limited to the tools used, such as questionnaires or psychological tests, but encompasses the overall data collection strat- egy (or design as detailed in the subsequent learning unit). It is important to differentiate

between data that is meant for analysis to draw conclusions. This distinction is between primary and secondary data: primary data being newly collected, and secondary data being existing data from a previous study. Secondary data can take the form of documen- tation, published papers, or previously collected data. Working with secondary data offers the advantage of bypassing the data collection step, thereby enabling immediate initia- tion of data analysis and evaluation. However, a disadvantage of using secondary data may arise when formulating new research questions, as such data may lack necessary information for this purpose. On the other hand, primary data collection, which includes case studies, observations, surveys, and experiments, carries its own set of advantages and disadvantages (Anderson, 2020). Commonly used tools in psychological research encompass questionnaires, psychological or neuropsychological tests, as well as biophy- siological and neuroimaging techniques.

##### Case Study

A noteworthy aspect of a case study is its comprehensive investigation of an individual, a small group of individuals, or an event (Anderson, 2020; Cuttler et al., 2019). Usually, these cases entail specific characteristics that are uncommon and not extensively understood (rare phenomena). Data collection tools may involve observation, interviews, or the administration of psychological tests or questionnaires. This method of data collection provides valuable in-depth insights into rare medical or psychiatric phenomena and serves as a valuable approach for formulating future hypotheses or research questions for larger-scale investigations. However, drawbacks include the absence of cause-and-effect relationships, the likelihood that the individual under examination is atypical, and a heavy reliance on the researcher’s subjective opinions and interpretations (Anderson, 2020; Cut- tler et al., 2019).

##### Observation (Naturalistic)

Observation of a behavior in its natural environment is often essential for researchers or clinicians seeking to understand behavior, particularly in terms of its nature, frequency, and contextual factors (Anderson, 2020; Cuttler et al., 2019). Observations can yield valua- ble insights in this regard. However, a key drawback associated with observation is the potential for individuals being observed to react to the observer’s presence, altering their behavior (reactivity). This phenomenon is known as the Hawthorne effect (Cuttler et al., 2019). It highlights the challenges of researching unaltered behaviors. Furthermore, estab- lishing cause-and-effect relationships in observations, similar to case studies, can be chal- lenging (Anderson, 2020).

##### Survey

This data collection method primarily involves the administration of questionnaires or tests, whenever feasible, to a sample derived from the population of interest, either online, over the phone, or in face-to-face settings (Anderson, 2020). It encompasses both quantitative and qualitative approaches, as it involves self-reports where participants directly express their thoughts, emotions, behaviors, and other relevant aspects of inter- est (Cuttler et al., 2019). Survey research can be susceptible to potential disadvantages related to the employed sample. If the sample is not truly representative of the popula-

tion, the resulting information can be inaccurate (Anderson, 2020). Additionally, biases may arise, such as interviewer and social desirability biases. In face-to-face or telephone surveys, respondents might not provide truthful answers due to potential interviewer biases or the expectations and opinions that may arise during the interaction (Anderson, 2020). Social desirability refers to participants’ inclination to respond in a manner they believe will be positively viewed (Cuttler et al., 2019). Participants may anticipate the sur- veyor’s intentions and exaggerate either positive or negative responses, driven by the per- ception of what the researcher would prefer their responses to be.

##### Experiment

This method entails creating a specific environment to manipulate IVs in order to assess their impact on the DVs, which are the elements being measured by the experimenter (Cuttler et al., 2019). Experiments are frequently utilized in psychological research and are regarded as the ideal method for investigating cause-and-effect relationships between variables. In an experiment, the researcher has the advantage of controlling irrelevant var- iables (extraneous variables) that could complicate the interpretation of findings (Ander- son, 2020; Cuttler et al., 2019). For instance, in an experiment aimed at studying concen- tration levels using a computerized test, a researcher may establish a controlled and quiet laboratory setting to prevent any disruptive noises during the experimental process. Despite the advantages, experiments are often at risk of human error and are not always feasible to conduct universally. Moreover, participants may be exposed to an unfamiliar and artificial setting that does not align with real-life experiences, potentially impacting their responses (Cuttler et al., 2019).

##### Correlational Studies

This method of data collection entails the measurement of variables in terms of their asso- ciations, or how strongly they are related to one another. Approaches such as surveys, observations, and experiments often examine the association between variables (Ander- son, 2020). Employing correlational approaches in data collection can help make predic- tions and sometimes extrapolate correlational findings from experiments to real-world contexts (Anderson, 2020). It is important to remember that a correlational relationship between variables does not imply causation.



**SUMMARY**

In psychological research, regardless of the specific project, the process generally involves several key steps. These include the initial formula- tion of research questions, followed by an assessment of their scientific and practical relevance. This involves determining whether the research question is scientifically testable and practically feasible. To gain an understanding of the current state of research on the chosen topic, it is important to review specialist published articles from scientific journals, as well as other relevant sources such as textbooks, specialist books, and preprints.

The next critical step involves formulating specific hypotheses based on the research questions. This is followed by careful consideration of the sample and the selection of an appropriate sampling method (probabil- ity or non-probability) to ensure the collection of data from a represen- tative sample of the population of interest. Ensuring that the sample mirrors the characteristics of the population is crucial for the study’s val- idity and generalizability.

When it comes to data analysis, researchers utilize primary (newly col- lected) or secondary (existing) data from previous investigations. Tech- niques for collecting data in psychological research may include sur- veys, observation, and experiments. Additionally, ethical considerations are paramount, and researchers must address and minimize any con- cerns related to participants and the handling of data.

The aim of research is to contribute new knowledge and information to the broader scientific community and society. Therefore, the findings and interpretations must be shared through the publication of scientific articles or in student undergraduate and postgraduate theses. This often takes the form of a comprehensive research report, which consists of specific sections, each responsible for conveying distinct information (abstract, introduction, methods, results, discussion, conclusion).

# UNIT 3

## RESEARCH DESIGNS

**STUDY GOALS**

On completion of this unit, you will be able to ...

* distinguish between experimental and other research designs.
* identify the available experimental research designs.
* name and describe available pre-, quasi-, and nonexperimental research designs.
* describe confounding variables are and define how they can be controlled in experi- mental planning.
* explain the essential conditions needed to draw causal conclusions in research.

### 3. RESEARCH DESIGNS

#### Introduction

Different research designs yield varying levels of internal and external validity in relation to the findings of an experimental situation. Internal validity pertains to the degree to which statistical results support a genuine association between a dependent variable (DV) and an independent variable (IV), and whether the results can be explained without the influence of other conditions apart from the manipulated IV. On the other hand, external validity refers to the degree to which results obtained from an experimental or research procedure can be generalized. In this learning unit, we will differentiate between experi- mental, quasi-, and nonexperimental research designs and their association with internal and external validity, both of which are essential in assessing causality between variables.

#### Experimental Research Designs

In an experimental research design, the experimenter investigates the variation of a DV under strictly defined conditions or characteristics of an IV. This process is referred to as an experiment, with the goal of addressing questions related to whether causality between variables genuinely exists, such as whether changes in the IV cause changes in the DV (Cuttler et al., 2019; McLeod, 2012)? In experiments, the level or condition of the IV is manipulated. Essentially, the experimenter systematically alters the levels of the IV so that different groups of participants experience varying levels of a variable, or the same group of participants experiences different levels of the IV at different times (Cuttler et al., 2019; Kantowitz et al., 2009; McLeod, 2012). For instance, a study investigates the potential association between consuming ice cream before bedtime and the frequency of night- mares among individuals. Participants are divided into two groups: one group eats ice cream before bedtime for a week, while the other group refrains from consuming ice cream. At the end of the period, the total number of nightmares reported by participants in each group is compared. The manipulation in this scenario is the ice cream condition, which creates two distinct conditions (ice cream vs. no ice cream) with different groups experiencing varying levels of the IV.

An alternative approach would involve measuring the same variable (nightmares) using the same set of participants. For example, the participants could consume ice cream before bedtime for one week and then refrain from consuming it for another week. In this scenario, the same participants would experience the different levels of the IV at different time points (Cuttler et al., 2019; Shaughnessy et al., 2012). What has been measured is the frequency of nightmares, which serves as the DV. To ensure that participants’ sleep would not be influenced by any external factors, such as noise or unexpected occurrences at home, all participants stay in the lab for an entire week, particularly during the nights. The objective is to exert as much control as possible over any other potential factors, apart from ice cream consumption, that could impact sleep quality. This control for additional variables, known as extraneous variables, beyond the IV being measured for the DV, is a

fundamental aspect of all experiments. It is essential for researchers to attempt to control, and thereby minimize variability, in the influence of any other variable apart from the IV on the DV (Cuttler et al., 2019; Kantowitz et al., 2009; McLeod, 2012; Shaughnessy et al., 2012).

As mentioned above, psychological research often involves experiments in which partici- pants are allocated to different groups or conditions (control and experimental) to investi- gate causal relationships between variables. This method of participant allocation in dif- ferent conditions is known as an experimental design, which includes three main types: between-subjects design, within-subjects design, and matched subjects design (Cuttler et al., 2019; Kantowitz et al., 2009; McLeod, 2017).

##### Experiments and Validity

After conducting an experiment, it is crucial to evaluate the validity of the specific findings (Cuttler et al., 2019). In psychological research, various types of validity are employed to assess the credibility of an experiment or research study. The primary types are internal validity, external validity, construct validity, and statistical validity (Cuttler et al., 2019; Kantowitz et al., 2009; Shaughnessy et al., 2012).

Internal validity refers to the relationship between the independent and the DVs, or in sim- pler terms, whether the IV is truly responsible for the changes observed in the DV (cause and effect). Empirical studies generally have high internal validity, when the manipulation of the IV leads to changes in the DV. Experiments tend to exhibit high internal validity as they are conducted under tightly controlled conditions, allowing for the manipulation of the IV by the experimenter and the control of confounding variables, thereby enabling causal conclusions to be drawn (Cuttler et al., 2019; Shaughnessy et al., 2012).

External validity refers to the generalizability of findings. When conducting a study or experiment using a sample from a population, the aim is to draw inferences about the broader population of interest. Research studies demonstrate high external validity when the individuals under investigation and the circumstances to which they are exposed are similar to those in the population being studied, often referred to as mundane realism (Cuttler et al., 2019). Experiments are often criticized for their external validity due to the controlled environment created by the experimenter, which some argue does not mirror real-life situations (e.g., when the IV is manipulated to cause changes in the DV) (Cuttler et al., 2019; Kantowitz et al., 2009; Shaughnessy et al., 2012).

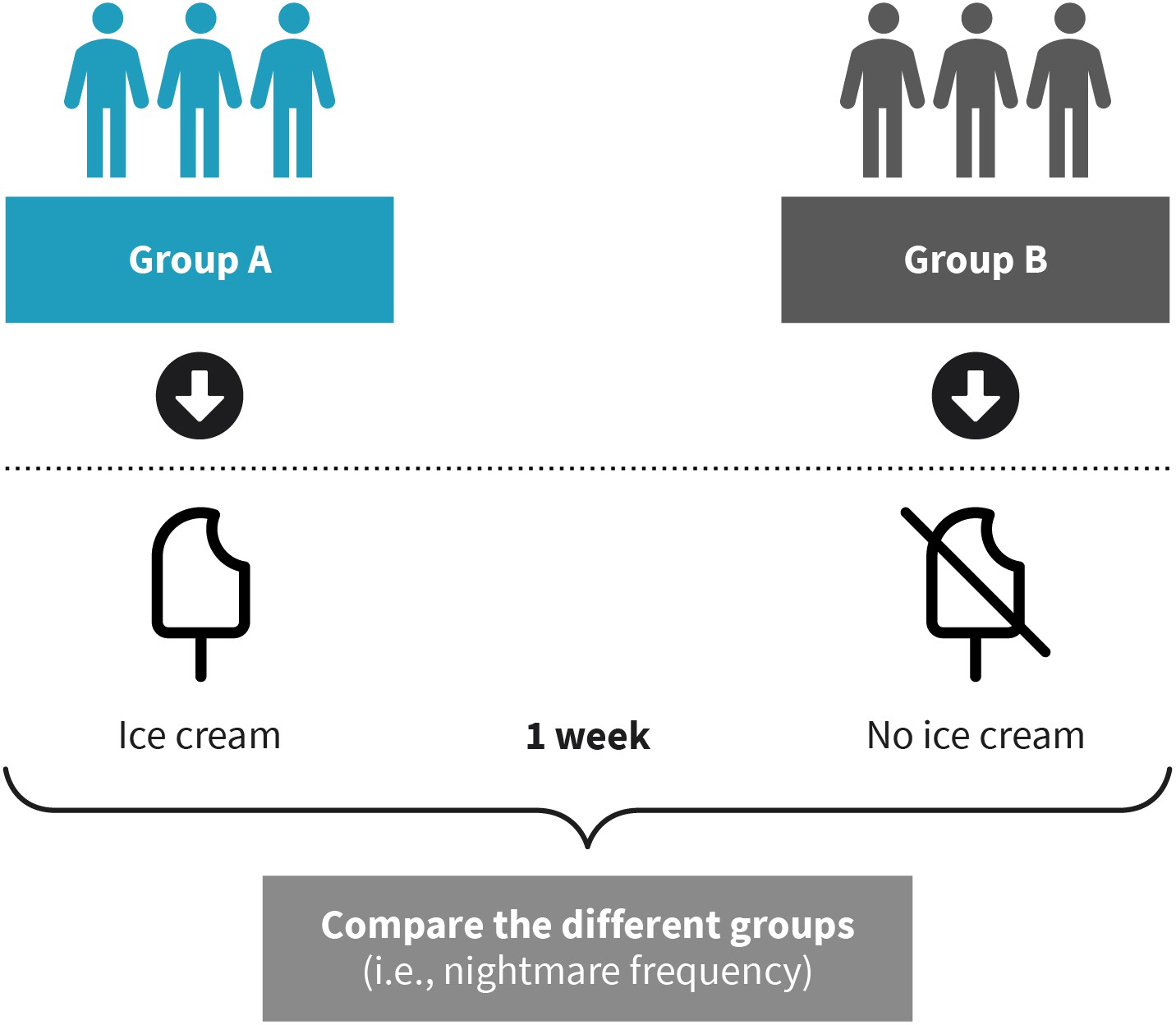
Construct validity refers to the degree to which an operational definition of a specific con- cept under study aligns with the actual measurement. For example, in a survey measuring satisfaction with a chocolate cake, the operational definition of satisfaction should encompass both appearance and taste. Failing to capture one of these elements would indicate low construct validity (Piretra, 2016).

Lastly, statistical validity relates to the appropriate statistical handling of data, ensuring that conclusions are drawn directly from statistical findings and that the proper statistical tests were employed (Cuttler et al., 2019; Shaughnessy et al., 2012).

##### Between-Subjects Design

This experimental design, also referred to as Unrelated Samples or Independent Samples design, involves each participant being assigned to only one condition of the IV (Cuttler et al., 2019; Shaughnessy et al., 2012). In this design, each participant is assigned to just one condition of the IV, resulting in two separate groups, each exposed to only one condi- tion. For example, if we have 40 participants in an ice-cream experiment, one group would exclusively consume ice cream before bedtime, while the other group would abstain from it. In this scenario, we would compare the frequency of nightmares between the two sepa- rate groups over the same period. The primary challenge with this type of research design is related to individual differences. When comparing two separate groups of participants, each consisting of different individuals, the increased likelihood of introducing systematic bias due to individual differences from one group to another is a concern. For instance, participants in one group may generally have good sleep quality, whereas those in the other group may have poorer sleep quality. This bias can significantly impact the findings. To mitigate this issue in between-subjects design, researchers utilize “randomization” (Shaughnessy et al., 2012). This involves the random allocation of participants to the two comparison groups, ensuring that each participant has an equal chance of being assigned to either group (McLeod, 2017).

Figure 6: Between-Subjects Design



Source: Evangelos Zois (2022).

###### Advantages and disadvantages

This type of design requires a larger sample size, resulting in a more time-consuming recruitment process compared to within-subjects design. Furthermore, individual differ- ences such as gender, ethnicity, socioeconomic background, and variations in life habits (e.g., sleep quality) can influence the findings. However, randomization can help mitigate these differences and reduce heterogeneity among comparison groups (Cuttler et al., 2019; McLeod, 2017). One advantage of a between-subjects design is that it eliminates practice or fatigue effects (order effects) because participants are exposed to only one condition (McLeod, 2017).

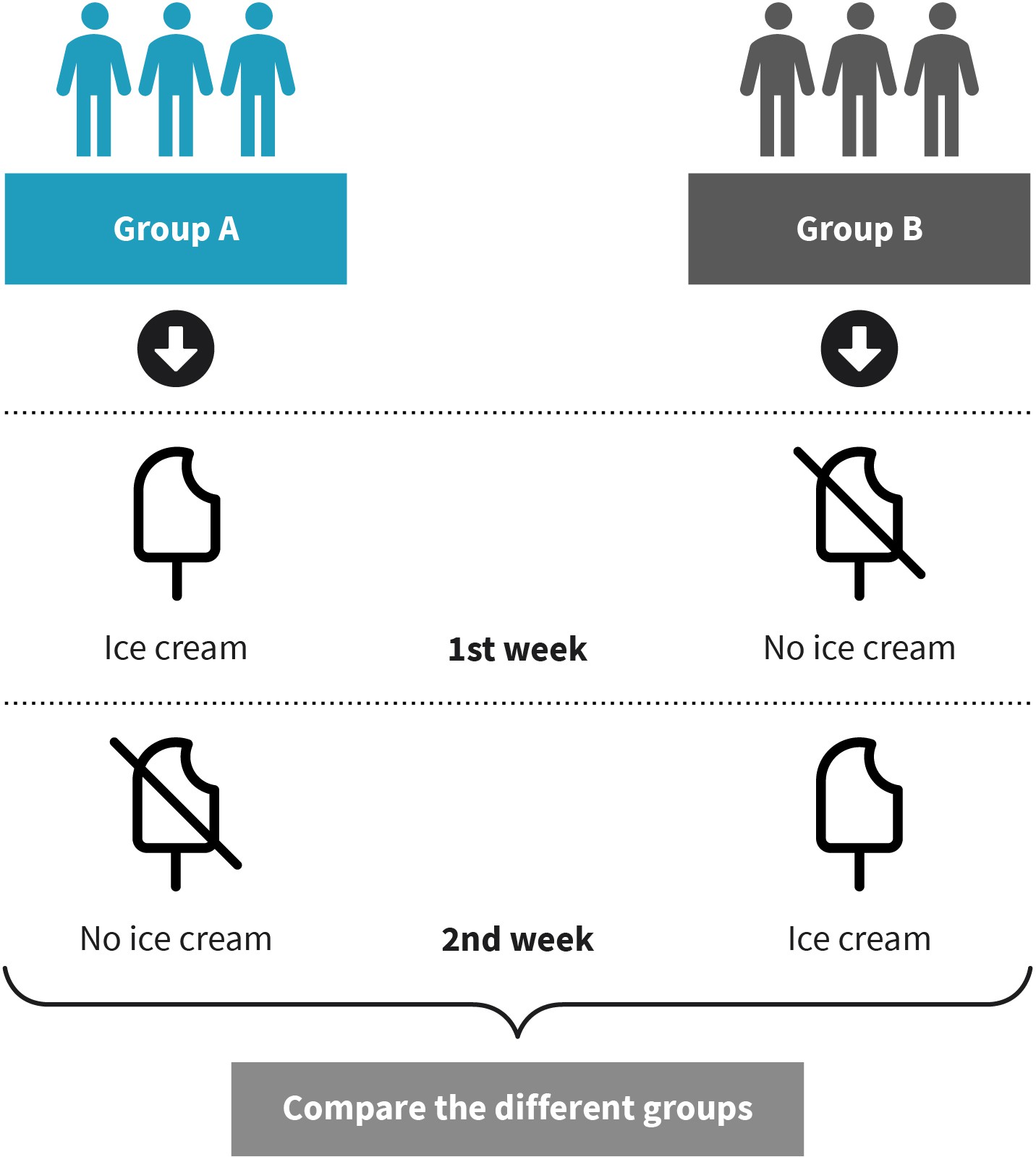
##### Within-Subjects Design

This type of experimental design, also known as Repeated Measures or Related Samples design, entails exposing participants to all conditions of the IV (Cuttler et al., 2019; Shaughnessy et al., 2012). In other words, participants are tested under all conditions involved in the experimental procedure (McLeod, 2017). Using the example of ice cream consumption and nightmares, when the same participants are tested under all conditions of the IV, such as consuming ice cream before bedtime for one week and then abstaining from it for another week, and the frequency of nightmares is measured and compared over the three-week period, this would be considered a within-subjects design.

Within-subjects design comes with a drawback known as “order effects,” which refers to the potential influence on participants’ behavior in a subsequent condition (such as refraining from eating ice cream in the second week) after being exposed to a preceding condition (for example, consuming ice cream in the first week) (Shaughnessy et al., 2012). In the ice cream example, the effects of consuming ice cream during the first week might influence the sleep patterns of participants in the following week, potentially leading to altered sleep patterns compared to their normal sleep quality.

Counterbalancing is a method used to address potential order effects in within-subjects designs. It involves testing different participants in different orders (McLeod, 2017; Shaughnessy et al., 2012). For example, in the ice cream experiment, 20 participants could be divided into two groups. One group would consume ice cream in the first week and abstain from it in the second week, while the other group would refrain from consuming ice cream in the first week and then consume it in the second week. By employing coun- terbalancing, the potential order effects are mitigated, and researchers can confidently attribute the findings to the conditions of the IV (ice cream vs. no ice cream) (McLeod, 2017).

Figure 7: Counterbalancing



Source: Evangelos Zois (2022).

###### Advantages and disadvantages

The primary benefits of a within-subjects design are the reduced participant variables introduced to an experiment, such as individual differences, as the same individuals are tested across each condition, and the requirement of a smaller sample size, making it more time-effective (McLeod, 2017). However, potential disadvantages of within-subjects design can stem from order effects, which may influence participants’ performance. For instance, participants anticipating the second condition may perform better due to prac- tice effects, or they may behave less effectively due to fatigue effects (Cuttler et al., 2019; McLeod, 2017; Shaughnessy et al., 2012).

##### Matched Subjects Design

In a matched subjects design, participants are paired, with each pair being matched based on key variables such as gender, age, socioeconomic status, ethnicity, or educational level (Cuttler et al., 2019; McLeod, 2017; Shaughnessy et al., 2012). Random assignment is then used to allocate one person from each pair to either the control or the experimental condi- tion (McLeod, 2017). For instance, in a study to assess the effectiveness of two exam prep- aration methods, i.e., the “standard” (control condition) or the “new” method (experimen- tal condition), a teacher might pair 50 students based on gender and age. These pairs of, for example, 14-year-old boys, 13-year-old girls, etc., are then randomly assigned to either the control or the experimental condition.

###### Advantages and disadvantages

A matched subjects design offers the advantage of reducing individual differences and eliminates the introduction of order effects, thereby negating the need for counterbalanc- ing (McLeod, 2017). However, some challenges associated with the matched subjects design include the difficulty of precisely pairing individuals, the time-consuming nature of achieving exact matches, and the potential loss of data from both individuals in a pair if one participant decides to drop out of the study (McLeod, 2017).

#### Controlling Confounding Variables

In psychological research, the experimenter is interested in examining a phenomenon, the DV, and observing how it changes due to another variable introduced and manipulated by the experimenter: the IV. It is common for the DV in a psychological experiment to be influ- enced by multiple variables other than the IV (Cuttler et al., 2019; Shaughnessy et al., 2012). These variables are called confounding variables, and they are not directly meas- ured but can still affect the cause-and-effect relationship between the IV and DV in a study (Thomas, 2020). When designing a research project, it is imperative to consider potential confounding variables and address them to ensure the credibility of the findings (Cuttler et al., 2019; Thomas, 2020). Confounding variables, also known as extraneous variables, are linked to both the IV and DV in a study. A variable is considered a confounder if it is correlated with the IV, often indicating a causal association, and has a causal relationship to the DV (Shaughnessy et al., 2012).

For instance, a researcher has gathered data on the correlation between sunburns and the consumption of cold juice. The findings of this study indicate that there is a connection between drinking cold juice and experiencing sunburns. Specifically, the more juice a per- son consumes, the greater their likelihood of getting sunburned. However, it would be incorrect to conclude that drinking cold juice directly causes sunburns. This study over- looks a confounding variable, namely temperature. Particularly during the summer, tem- peratures tend to be high. In such conditions, individuals naturally seek ways to cool down, such as consuming cold beverages like juice. Additionally, people spend a signifi- cant amount of time outdoors during the summer, exposing themselves to the sun, which is the actual cause of sunburns (Thomas, 2020). Confounding variables play a crucial role

in psychological research as they ensure the internal validity of a study. Internal validity refers to the assurance that any changes observed in the DV are a result of the experimen- ter’s manipulation of the IV, rather than being influenced by other variables or factors (Cut- tler et al., 2019).

There are methods to reduce the impact of confounding variables in a research project (Thomas, 2020). Randomization is a practical approach to ensure each element of the sample has an equal likelihood of being included in any of the experimental or control conditions (LaMorte, 2021). Especially in a large sample, any unidentified confounding variables may be shared between the control and experimental groups, leading to equal distribution of confounds among the groups. Another method involves restricting study inclusion to subjects with only one known confounding variable in a research study or experiment (LaMorte, 2021). For example, when gender is a confounding variable, restrict- ing inclusion to only men or women may be necessary. Similarly, when age is a confound- ing variable, inclusion can be restricted to certain age groups. However, it is important to consider that restrictions may limit the generalizability of findings to the excluded groups (LaMorte, 2021; Thomas, 2020). Another method to limit the impact of confounding fac- tors, particularly those that are well-known, is to recruit and match participants according to the confounding variables. For instance, when studying brain morphology and struc- ture, age is an important factor. If the individual is too young, the brain may not have fully developed yet, while if they are older, brain morphology may have already started to dete- riorate. The optimal approach is to enlist individuals belonging to the same age group. Although matching can be both time-consuming and costly, it is valuable in accounting for extraneous variables that are difficult to measure (LaMorte, 2021).

#### Pre- and Quasi-Experimental Research Designs

Researchers frequently employ specific research designs when it is difficult to create experiments. They typically opt for these methods to gain insight into cause-and-effect relationships in real-world scenarios where it is challenging to create optimal experimen- tal conditions.

##### Pre-Experimental Research Designs

When it is not possible to conduct a true experiment, an experimenter may decide to ini- tiate a pre-experimental design. Pre-experimental designs are typically implemented prior to conducting a true experiment. They are commonly used when testing a new interven- tion, often with relatively small sample sizes. For instance, a research team may be inter- ested in seeking funding to support their novel intervention idea. However, before con- ducting a large-scale experiment, they need to provide initial evidence of its effectiveness through a pre-experimental design (Cuttler et al., 2019; Kantowitz et al., 2009; Shaugh- nessy et al., 2012). Pre-experimental designs encompass various types, which are descri- bed below.

###### One-group posttest only design

In this design, an intervention or treatment is implemented, and the DV (improvement resulting from the intervention) is measured only after the intervention has been com- pleted (Cuttler et al., 2019; Shaughnessy et al., 2012). For instance, consider a scenario where a researcher is interested in examining the effects of a specific relaxation technique on a student population before their final exams. The relaxation technique is introduced and practiced for a week, and relaxation levels are measured immediately afterward. However, this design lacks a control group, making it difficult to determine the baseline relaxation levels of the students prior to the intervention. It is possible that the students were already experiencing high levels of relaxation before the intervention.

###### One-group pretest-posttest design

In this design, the experimenter first measures the DV before implementing the treatment, and then measures it again after the treatment has been implemented (Cuttler et al., 2019). Referring back to the previous example of the relaxation technique intervention for students before their exams, the experimenter would initially measure relaxation levels in the target population. Then, the relaxation technique would be implemented for one week, followed by another measurement of relaxation levels in the students. This design resembles a within-subjects design, where the same group of participants is tested under all conditions. However, it is important to note that there is no counterbalancing involved in this design. If the post-intervention relaxation levels are higher or better than the pre-intervention levels, it can be tentatively concluded that the intervention was effective. However, it is not entirely accurate to attribute the observed change or improvement solely to the intervention, as there may be other factors or explanations that could threaten the internal validity of the findings (Cuttler et al., 2019; Shaughnessy et al., 2012).

The following potential alternative explanations pose a threat to internal validity (Cuttler et al., 2019):

* **history:** It is possible that events or factors occurring between the pretest and posttest may be responsible for the observed change (Howitt & Cramer, 2017; Cuttler et al., 2019). For instance, as finals approach, the weather may improve, leading students to engage in activities such as swimming, which could contribute to maintaining high relaxation levels during both the pretest and posttest periods.
* **maturation:** It is also possible that the students naturally changed between the pretest and posttest periods, independent from the intervention (Howitt & Cramer, 2017; Cut- tler et al., 2019; Shaughnessy et al., 2012). This change could be attributed to their pre- vious experiences with exams, leading to psychological maturation and improved prep- aration. This is particularly relevant if the time gap between the testing periods exceeds a week.
* **testing:** This relates to the phenomenon of measuring a response before an interven- tion, and how this may influence responses during the subsequent measurement (Howitt & Cramer, 2017; Cuttler et al., 2019). For instance, if students complete a ques- tionnaire to assess their relaxation levels before the intervention, this initial measure- ment might influence their responses to the same questions during the posttest. It is possible that the act of completing the pretest questionnaire made the students more

aware of their anxiety levels, enabling them to manage and control their anxiety, thereby maintaining low anxiety levels and high relaxation levels, regardless of the intervention.

* + **instrumentation:** This is another potential threat to the internal validity of a study employing a one-group pretest-posttest design (Howitt & Cramer, 2017; Cuttler et al., 2019). It relates to the factor of familiarity with the measurement procedure or tool used to assess relaxation. For instance, students may have taken the pretest measurement very seriously simply because it was a new experience, and they were unfamiliar with it. However, during the posttest, they might approach the measurement differently, feeling bored or less interested and attentive (Cuttler et al., 2019).
  + **regression to the mean:** This phenomenon is related to a statistical concept which sug- gests that participants who provide a very high score on one testing occasion are more likely to provide a lower score on a subsequent testing occasion (Cuttler et al., 2019; Shaughnessy et al., 2012). For instance, consider a scenario where the students selected to receive the relaxation intervention were those who initially scored particularly low on the relaxation scale, indicating high levels of tension. During the posttest, these stu- dents may exhibit reduced tension, not necessarily due to the intervention, but as a result of regression to the mean.
  + **spontaneous remission:** This phenomenon is observed in various medical and psycho- logical conditions, where improvement occurs over time without any specific treatment or intervention. For instance, in depression, research has indicated that up to 53% of untreated cases experience remission within 12 months (Whiteford et al., 2013).

It is crucial to consider all potential threats to internal validity and to exercise caution when interpreting findings from pretest-posttest designs. One effective approach to miti- gate the risk to internal validity is to incorporate a control group that does not receive the intervention (Cuttler et al., 2019; Shaughnessy et al., 2012).

###### Interrupted time series design

It describes an experimental design involving pretest and posttest measurements at sev- eral intervals over a period (Cuttler et al., 2019). For instance, a company manager aims to assess productivity over the course of a year, with measurements taken weekly. In this sce- nario, the term “interrupted” in interrupted time series refers to the point at which the intervention is introduced. In the given example, the intervention involved reducing work- ing hours from 12 to 9. The manager observed an increase in productivity after the reduc- tion in working hours, which remained high in the subsequent months. Consequently, the manager attributed this change to the reduction in working hours. This design differs from a standard pretest-posttest design in that it incorporates multiple pretest-posttest meas- urements within a specified period (Cuttler et al., 2019; Shaughnessy et al., 2012).

##### Quasi-Experimental Research Designs

**quasi** meaning “in parts,” “seemingly,” “almost”

A **quasi**-experimental design is similar to an experiment (Cuttler et al., 2019; Shaughnessy et al., 2012). The main difference is the absence of randomization, specifically the random assignment of research participants to either the control or experimental conditions or groups (Cuttler et al., 2019). In quasi-experimental research, a comparison group, similar to a control group, is often utilized, but without random assignment and the manipulation

of the IV or, in other words, predetermined conditions (Kantowitz et al., 2009). The IVs of interest, such as gender or diagnosis, are naturally occurring, and a quasi-experimental research study may be designed to examine differences between these variables, as well as race or ethnic group (Kantowitz et al., 2009; Shaughnessy et al., 2012). Due to absence of random allocation, the compared groups might be dissimilar in certain ways (not equiv- alent), leading to a between-subjects design with nonequivalent comparison groups. This is also known as “nonequivalent comparison groups design,” and it resembles a true experiment. However, random participant allocation is not employed, and the comparison groups may already exist with inherent dissimilarities (Cuttler et al., 2019). In summary, it can be concluded that IVs in quasi-experimental research occur naturally and involve a subject variable (the IV) (Kantowitz et al., 2009; Shaughnessy et al., 2012).

In clinical studies examining memory ability following a diagnosis of depression, one group consists of individuals diagnosed with depression, while the comparison or control group comprises individuals without a diagnosis of depression. If the experimenter has effectively controlled for other potential confounding factors that could influence memory function, any differences observed between the two groups would be attributed to the diagnosis. In this scenario, there was no random allocation to either group, and the experimenter did not create or manipulate the IV and its levels, as they preexisted (Cuttler et al., 2019). It is important to note that the comparison groups may not be as similar as desired. In this example, the IV is the diagnosis (depression vs. non-depression), and the DV is memory function.

##### Types of Nonequivalent Groups Designs

Nonequivalent groups designs are research designs used when random assignment is not possible, allowing researchers to compare the effects of a treatment or intervention on different groups. Some common types of nonequivalent groups designs will be introduced below.

###### Posttest-only nonequivalent groups design

In this design, one group of participants is exposed to a treatment or intervention, while a nonequivalent group of participants does not receive the treatment or intervention. The two groups are then compared at the end (posttest) (Cuttler et al., 2019). For instance, consider a study evaluating the effectiveness of a new psychological intervention for fear of flying. Two groups of depression patients are involved; one group receives the interven- tion, while the nonequivalent group does not. The comparison is then made to determine if individuals who received the intervention show improved symptoms related to their phobias. To enhance internal validity and mitigate the effects of confounding variables, the experimenter can take measures such as matching participants in both groups based on age, gender, socioeconomic background, duration of diagnosis, and other relevant fac- tors. However, without true randomization, there is a relatively high likelihood of other confounding variables influencing the results (Cuttler et al., 2019; Shaughnessy et al., 2012).

###### Pretest-posttest nonequivalent groups design

This design represents an enhanced version of a nonequivalent groups design, incorporat- ing a pretest condition. The treatment group (experimental group) undergoes a pretest, followed by the intervention or treatment, and then a posttest (Cuttler et al., 2019). Simi- larly, the nonequivalent control group experiences a pretest but does not receive the inter- vention or treatment, and then undergoes a posttest. The experimenter aims to determine not only whether participants exposed to the intervention exhibit improvement, but also whether their improvement surpasses that of the group not receiving the intervention (Cuttler et al., 2019). For instance, if patients diagnosed with anxiety disorder who receive the intervention demonstrate greater symptom improvement compared to those who did not receive the intervention, it is likely attributed to the intervention. However, it is impor- tant to note that this design type does not entirely eliminate the presence of confounding variables (Cuttler et al., 2019; Shaughnessy et al., 2012).

###### Interrupted time-series design with nonequivalent groups

This quasi-experimental design involves collecting data at various time points (intervals) before and after the treatment in two or more nonequivalent groups (Cuttler et al., 2019; Shaughnessy et al., 2012). It enables the examination of the effect of a treatment or inter- vention over time. For instance, consider a scenario where a product is sold at a specific price, and the interest lies in the sales associated with this product over a 12-week period. It is observed that sales remained steady for the first 4 weeks after the product became available, but from week 5 onwards, sales increased following a decision by the company to decrease the price. Data collected each week until week 12 revealed a consistent increase in sales from the 5th week onward. However, a drawback associated with this design type pertains to circumstances or history, as unexpected events beyond the experi- menter’s control may influence the cause-and-effect relationship between the independ- ent and dependent variables. Additionally, maturation, which refers to natural processes occurring over time, could also impact the results, such as an increase in sales for the selected product regardless of the price change.

###### Pretest-posttest design with switching replication

In this scenario, a pretest-posttest design with a switching replication design is employed. This involves administering a pretest of the DV to nonequivalent groups, then exposing one group to a treatment while a nonequivalent control group does not receive the treat- ment. The DV is measured again, and subsequently, the treatment is added to the control group. Finally, the DV is measured once more (Cuttler et al., 2019). For example, consider an experimenter proposing a physical activity intervention for individuals with depres- sion. The sample consists of depression patients in one group and a nonequivalent control group of students experiencing depression (Howitt & Cramer, 2017; Cuttler et al., 2019; Shaughnessy et al., 2012). In this design, the first step involves measuring depression in both groups, followed by introducing the intervention to the depression patient group. Depression levels are then measured in both groups. If the intervention is effective, a decrease in depression symptoms is expected in the patient group exposed to the physical activity intervention, while no change is anticipated in the group of students with depres- sion (no exposure to the intervention). Subsequently, the group of students with depres-

sion is exposed to the same intervention, while the patient group remains exposed to the intervention. It is now anticipated that depressed students would also show a decline in depression symptomatology (Cuttler et al., 2019). This design leverages replication, allow- ing the effects of the intervention to be observed in one group and then replicated in the second group. Additionally, history effects are considered, as it is unlikely that an event potentially influencing the findings would occur exactly at the same time as the interven- tion introduction in both groups, given that the intervention is not introduced simultane- ously in both groups (Cuttler et al., 2019; Shaughnessy et al., 2012).

###### Switching replication with treatment removal design

This design type differs from the previous design described above. Instead of utilizing a pretest-posttest design with switching replication, in which one group receives the inter- vention while the other group is exposed to it at a later time while the first group contin- ues to receive treatment, the switching replication with treatment removal design involves removing the intervention from the first group when it is introduced to the sec- ond group (Cuttler et al., 2019; Shaughnessy et al., 2012). For instance, in the previous depression example, patients with depression receive the physical exercise intervention initially, while the students with depression do not. After a week, depression levels are measured in both groups. If the intervention is effective, depression should decrease in the patient group but not in the other group. Subsequently, the intervention is removed from the patient group, and at the same time, the students are instructed to engage in physical activity. After a week, depression levels are measured in both groups again. If the treatment had an effect, the expected outcome would be a decrease in depression levels in the student group and an increase in the patient group (as they no longer receive the intervention) (Cuttler et al., 2019; Shaughnessy et al., 2012). This design provides valuable evidence on the efficacy of the intervention.

#### Nonexperimental Research Designs

Nonexperimental research is defined as the type of research that does not involve the manipulation of an IV. Instead, an experimenter conducting nonexperimental research examines variables as they naturally occur in settings such as a lab or in nature (Cuttler et al., 2019). While an experimenter still measures a DV, they cannot attribute differences observed in the DV to changes in the IV (Cuttler et al., 2019). The choice of research approach depends on the circumstances and interests of the experimenter. Nonexperi- mental research designs are utilized when it is not feasible to manipulate an IV, when the primary focus of the research is not to examine cause-and-effect relationships, when the research is solely focused on one variable, and when the purpose of the research is purely exploratory (Cuttler et al., 2019).

Nonexperimental research encompasses three main types: cross-sectional, longitudinal, and correlational research. Each of these types will be described below.

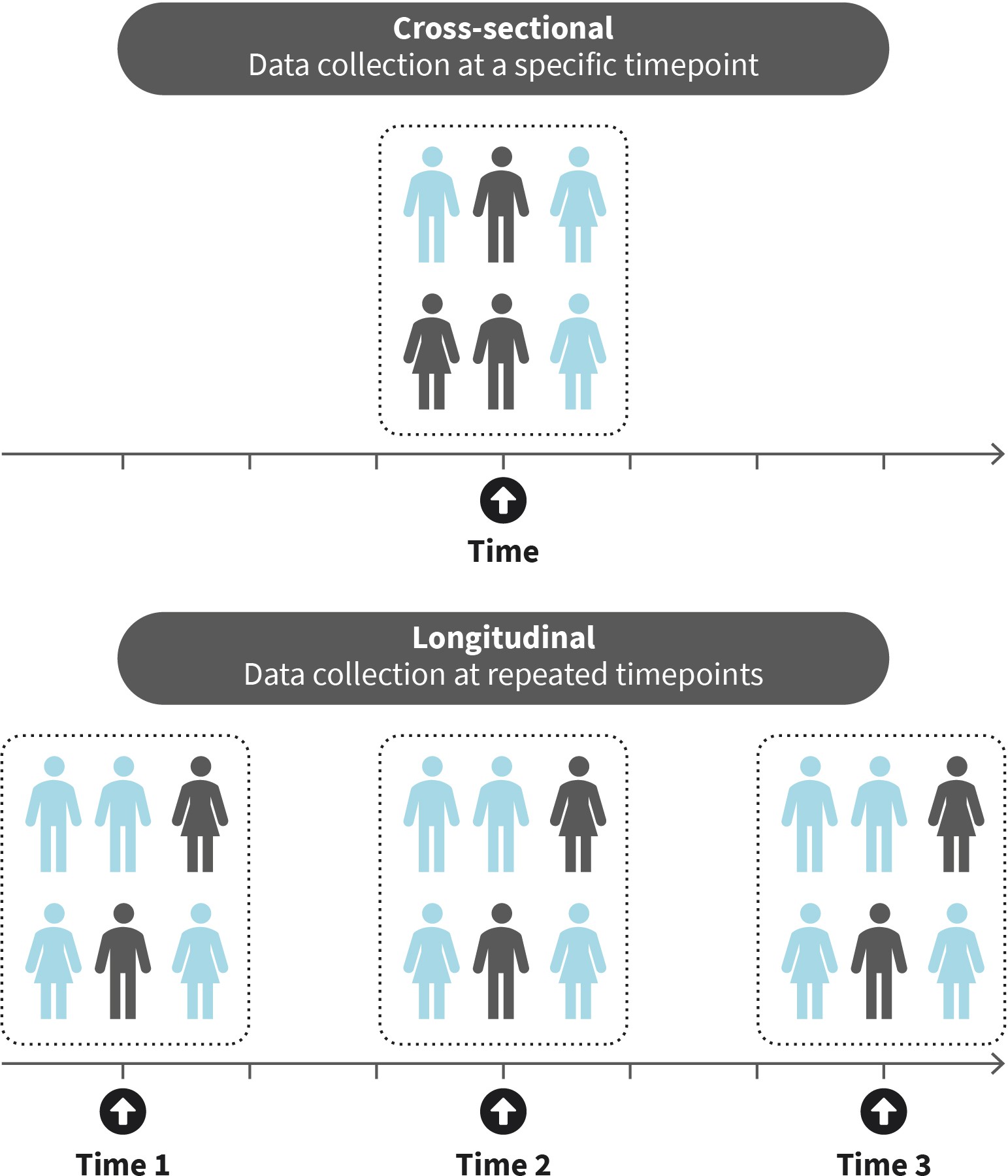
##### Cross-Sectional Research Design

This research design involves collecting data from different participants at a specific moment, without influencing any of the variables being observed (Howitt & Cramer, 2017; Cuttler et al., 2019; Thomas, 2021). For instance, a researcher may be interested in gather- ing data on alcohol addiction from a large sample to characterize the disease in terms of emotional, neuropsychological, neurobiological, and genetic variables. In this example, data is collected at specific moments in time for many different participants. A cross-sec- tional design is suitable when the aim of the study is to investigate prevalence rates and general characteristics of a particular outcome (in this case, alcoholism) at a specific point of time (Thomas, 2021). Researchers often opt for cross-sectional studies due to their cost and time efficiency, and they are frequently used as an initial step to gather a substantial amount of data that can serve as the foundation for more complex research (Howitt & Cramer, 2017; Cuttler et al., 2019; Thomas, 2021).

##### Longitudinal Research Design

Unlike cross-sectional research, longitudinal research design involves gathering data from the same participants repeatedly over time. An example of longitudinal research is the IMAGEN study, a longitudinal research project conducted across multiple research centers in Europe. Healthy adolescent participants were recruited at the age of 14 and followed up at the ages of 16, 19, and 22. The study aimed to understand human brain development and identify vulnerability profiles in individuals who may later develop addiction and other psychiatric diseases (Mascarell Maričić et al., 2020; Schumann et al., 2010). Longitu- dinal studies provide valuable insights into changes over time within the same individuals, using consistent measures and conditions at each time point. In contrast, a cross-sectional study would involve collecting data at a single time point, such as at age 14 or 16. Both research approaches have their own utility depending on the research question. While cross-sectional research is cost-effective and allows for the collection of a large amount of data for comparisons between different participant groups, it may not provide informa- tion on cause-and-effect relationships or long-term trends (Thomas, 2021).

Figure 8: Cross-Sectional vs. Longitudinal Research Design



Source: Evangelos Zois (2022), based on Thomas (2021).

##### Correlational Research Design

This type of nonexperimental research design involves the examination or measurement of two variables to investigate a statistical relationship, using the statistical technique of correlation. In this type of nonexperimental research, there is no control of extraneous var- iables (Cuttler et al., 2019). Researchers may use correlational research to explore poten- tial causal relationships between two variables, or they may simply be interested in describing the strength of the correlation between the variables. If the correlation is strong enough, it becomes possible to predict the scores of one variable based on the scores of the other, using a statistical technique called regression (Cuttler et al., 2019). Cor-

relational research design is also valuable for constructing new psychological tools, such as questionnaires, to measure specific variables. Through correlation, researchers can assess the validity and reliability of these tools (Cuttler et al., 2019).

**Causality** describes the relationship between cause and effect

#### Requirements for Causal Conclusions

In psychological research, there is often an interest in examining **causality,** such as under- standing the influence of one variable on another. For example, if a traumatic event leads to emotion dysregulation and subsequently to depression. However, drawing conclusions about causality is not feasible in nonexperimental studies because the effects of a third variable cannot be eliminated. Longitudinal studies, quasi-experiments, and pre-experi- ments can only suggest causality over time, and randomization with a relatively large sample is necessary to systematically exclude the effects of third factors. In order to draw causal conclusions between two variables, the following conditions should be met:

##### Covariation

This concept relates to the IV and the DV, specifically emphasizing that changes in the IV must be accompanied by corresponding changes in the DV. In other words, the relation- ship between the two variables must demonstrate that as one variable increases or decreases (in terms of numerical values), the subsequent values of the second variable also exhibit an increase or decrease. If variable A influences variable B, the two are said to be associated or to covary (Trivedi, 2020). For instance, if vaccination against COVID-19 provides protection, it would be expected that individuals who have been vaccinated against COVID-19 would be healthier during a COVID-19 outbreak compared to those who have not been vaccinated (Trivedi, 2020).

##### Temporal precedence

The IV should precede the DV, meaning that the cause should occur before the effect. This condition is not contingent on when the variables were measured (Trivedi, 2020). For instance, there is a correlation between economic growth and an increase in individuals’ income within society. It would be illogical to observe an improvement in income before an economic growth (Trivedi, 2020).

##### Exclusion of alternative explanations

It is essential to eliminate the influence of significant third variables (Trivedi, 2020). As pre- viously mentioned, these additional variables are also known as confounding or extrane- ous variables. They impact the relationship between the IV and the DV, and although it may initially appear that changes in the IV cause changes in the DV, this is not necessarily the case. In correlational research designs, the researcher lacks control over confounding variables. For instance, a correlation was identified between individuals who owned expensive cars and longer lifespans (Trivedi, 2020). Essentially, owners of expensive cars lived longer. Can we confidently assert that owning an expensive car increases lifespan based on this correlation? The answer is no. There are certainly other factors at play. A

contributing factor might be wealth, as those who are affluent can afford expensive cars. Overall quality of life is better for the wealthy (better life insurance, healthier eating habits, more vacations, etc.), which could better explain the association between owning an expensive car and longevity (Trivedi, 2020).



**SUMMARY**

The selection of a research design is a crucial phase in the research proc- ess, as different research designs have implications for internal and external validity. In an experimental research design, participants are randomly assigned to either the experimental or control group or condi- tion. The groups or conditions created represent the independent varia- ble (IV), which is manipulated by the researcher. The primary experi- mental research designs discussed in this unit are the between-subjects design, the within-subjects design, and the matched-subjects design. Randomization is essential in all types of experimental designs. How- ever, randomization is not present in pre- and quasi-experimental research designs. Pre-experimental designs involve either a pre-test- post-test schedule of a group or the post-comparison of a newly recruited group of participants with results from previous investigations. Quasi-experimental designs follow a similar approach and analytical procedure to experimental designs, with the only difference being that group assignment is not random (i.e., there is no randomization). Nonex- perimental designs include cross-sectional, longitudinal, and correla- tional designs that do not allow researchers to manipulate the IV. In principle, statements about causality can only be confidently made in the case of randomization. Key criteria for establishing causal conclu- sions include the covariation between the independent and dependent variables, the temporal precedence of the IV, and the elimination of alternative explanations.

# UNIT 4

## QUANTITATIVE METHODS

**STUDY GOALS**

On completion of this unit, you will be able to ...

* describe the different quantitative survey methods available.
* explain how psychological questionnaires and tests are designed.
* describe what measurement theory and test theory entail.
* indicate how biases in quantitative surveys can be avoided.
* list the quality criteria that define quantitative survey methods.

### 4. QUANTITATIVE METHODS

***2024-06-12 11:39:37***

--------------------------------------------

No referring to IU in text. Also, DE reference

#### Introduction

Consider the following scenario: A student at IU aims to investigate, as part of their thesis project, the stress experienced by refugees in Germany. After conducting an extensive liter- ature review, the student discovers that previous research has explored the topic of stress experienced by refugees in Germany, but there is limited literature available on the stress levels of unaccompanied refugee children. Additionally, the student finds that previous studies utilize various quantitative survey methods to investigate stress. For instance, some studies employ questionnaires to assess everyday stress or stress in specific situa- tions (e.g., social), while others use skin conductance responses as an indicator of stress during stress-inducing tasks. Furthermore, some researchers measure cortisol concentra- tion in hair as a marker for chronic stress levels. The student recognizes the significance of the data collection approach and the need for careful consideration. This unit offers an overview of quantitative survey methods.

#### Characteristics of Psychological Methods

The objective of quantitative survey methods is to capture phenomena relevant to psy- chology and convert them into measurable forms, such as nominal, ordinal, or interval- scaled characteristics. Researchers aim to describe phenomena as accurately as possible. Psychological research can take various forms. In some cases, psychologists observe behavior, while in others, research primarily focuses on collecting data on conscious expe- riences, thinking processes, or personality traits (McCombes, 2019). Research psycholo- gists often have an interest in unconscious processes and their manifestations. To accu- rately record unconscious processes, which are often unclear even to the participants themselves, researchers can translate abstract ideas, thoughts, and notions into quantifia- ble observations (operationalization) by employing appropriate data collection tools (Cut- tler et al., 2019; McCombes, 2019).

Psychological research frequently encounters the issue of reactivity, which pertains to biases introduced to the data as a result of participants being aware that they are under investigation (Cuttler et al., 2019). An established example of reactivity is the Hawthorne effect, which denotes a form of reactivity in which individuals (or participants) modify their behavior from their typical patterns as a result of being aware that they are under investigation (Cuttler et al., 2019; Shaughnessy et al., 2012).

Issues arising from reactivity can be mitigated through strategies implemented by researchers based on the objectives or procedures involved in a research investigation. These strategies aim to measure or test characteristics that participants cannot influence. For instance, depending on the research objectives, investigators often withhold informa-

tion about the study from test participants. However, it is crucial to prioritize ethical con- siderations in such cases. In experimental psychological research, this practice is known as deception, which, although generally discouraged, is sometimes considered ethically appropriate. When deception is incorporated into the research methodology, it is essential for participants who have been deceived to be fully informed about the actual aims of the research they participated in, but only after the completion of their involvement. It is important to note that deception never involves treating people disrespectfully, and the rights of participants regarding confidentiality, anonymity, and minimal risk are fully respected regardless of the study’s objectives. Further discussion on this topic will be pro- vided in one of the upcoming units.

Another approach to address reactivity is to utilize implicit versus explicit instruments and tools for data collection (Barakji, 2017). Implicit methods indirectly capture unconscious processes based on behavioral aspects, while explicit methods directly examine individu- als’ attitudes. Implicit methods are cognitive measurement tools that can capture auto- matic psychological characteristics that participants may be either unwilling or unable to share, as they might not be fully aware of the processes involved (Barakji, 2017). For instance, individuals may consciously believe that gender should be equally associated with science as a profession, but their automatic response (implicit) may indicate that most people associate science more with men than with women.

A classic example of a tool used to measure implicit processes is the “Implicit Association Test” (Greenwald et al., 1998). This technique is based on the observation that research participants tend to respond similarly when presented with **stimuli** representing related concepts (e.g., “positive” and “flower”) as opposed to when unrelated stimuli are pre- sented (e.g., “positive” and “gun”) (De Houwer, 2001). The test was employed in a 1998 study on prejudice between Bavarian and Northern Germans. The research team discov- ered that Northern Germans rated typical Bavarian names negatively, just as Bavarians rated Northern German names negatively. These implicit measures of prejudice were cor- related with information from explicit measurement procedures in which participants were directly asked about their prejudices (Neumann et al., 1998).

#### Introduction to Measurement Theory

The fundamental concept of measurement theory pertains to the relationship between the characteristics being measured and the tools used for measurement. This presents a significant challenge in quantitative psychological research: how to quantify psychologi- cal characteristics in numerical form (Michell, 1997). Physical quantities are routinely measured in everyday life using known units that can be expressed as numerical values (e.g., temperature in Celsius degrees, volume in liters or cubic centimeters, weight in kilo- grams, etc.). In psychology, the definition of a unit is not as clearly defined, and the meas- urement based on units cannot be directly applied to psychological variables (Döring & Bortz, 2016). Measurement theory addresses the definition and rules for transferring measurements, and it was introduced to psychology by Stevens (1946). According to Ste- vens, any quantitative measurement serves as a guide for observed or empirical elements and the assignment of numerical values to those elements (Stevens, 1935; 1946). His pre-

**Stimuli**

something that evokes a reaction (e.g., images, words)

cise definition of measurement was “the assignment of numerals to objects or events according to rules” (Stevens, 1946, p. 677). Therefore, measurement involves assigning numbers to objects or events, or constructs based on established rules in such a way that numerical relations correspond to empirical ones (Michell, 1997).

The following terminology is relevant to measurement theory (Döring & Bortz, 2016):

* + - **relation**: It describes the relationship between objects (persons, events, etc.) according to specific relational rules. For example, in a sample of five people, a relation rule such as “same gender” can be specified (Döring & Bortz, 2016).
    - **empirically relevant**: This is a set of objects or events that define a relation. The term “empirically relevant” always refers to the characteristics, properties, or features of the object. For example, determining the height of three individuals based on their appear- ance: Julia is taller than Sabine and Laura. In this instance, a set of objects is arranged according to height (empirically relevant).
    - **numerically relevant**: A set of numbers (e.g., 4, 3, 5, 6) can be related using mathemati- cal functions (>, <, = etc.). For instance, one can determine the height of three people: Sabine = 172 cm, Julia = 170 cm, and Anna = 168 cm. This means that, according to height, Sabine > Julia > Anna, and Anna < Julia < Sabine. In this example, a set of num- bers is sorted according to size (numerically relevant).
    - **homomorphism**: It ensures that an object’s actual properties are clearly assigned numerical values. Each empirical element is given exactly one number (numerically rel- evant). For instance, the gender of test participants can be recorded, and the numerical value of 3 can be assigned to each male and 4 to each female participant. Any other numerical assignment is possible as long as it leads to a clear distinction between gen- ders. However, the numeric value of 3 indicates a male participant, but it does not spec- ify the participant’s identity.

Table 3: Empirical Characteristics (A) and Numerical Definition (B)

|  |  |  |  |
| --- | --- | --- | --- |
| **A** | **Name** | **Gender** | **Education level** |
| **1** | Nigel | Male | Bachelor |
| **2** | Anna | Female | Master |
| **3** | Benny | Male | Bachelor |
| **4** | Stuart | Male | Doctorate |
| **5** | Maria | Female | Master |
| **B** | **Name** | **Gender** | **Education level** |
| **1** | Nigel | 3 | 2 |
| **2** | Anna | 4 | 3 |
| **3** | Benny | 3 | 2 |
| **4** | Stuart | 3 | 4 |

|  |  |  |  |
| --- | --- | --- | --- |
| **5** | Maria | 4 | 3 |

Source: Evangelos Zois (2022).

#### Questionnaires and Ratings

A survey is a data collection method that incorporates both qualitative and quantitative approaches to investigate variables of interest. It involves the use of self-reports, such as questionnaires and interviews, to gather information (Howitt & Cramer, 2017; Cuttler et al., 2019). Surveys are typically standardized and validated, meaning that the answer options are precisely defined. Participants, or survey responders, are asked to provide direct information pertaining to their thoughts, behaviors, and attitudes (Cuttler et al., 2019). Therefore, when selecting a questionnaire tool, researchers must carefully consider the characteristics or data related to the variables of interest, the theoretical concepts underlying the chosen tool, and the psychometric properties of the questionnaire, such as reliability and validity.

Many questionnaires are protected by copyright, which means that individuals who wish to use them must obtain formal permission and pay any required fees. Other question- naire tools may only require formal permission from the developers or authors, while some tools may not require any permission at all. The status of each tool, depending on its origin, is always indicated for researchers conducting survey research projects and utiliz- ing questionnaires for data collection. Questionnaires may exist in multiple languages, depending on their origin (Byrne et al., 1994; Giannakou et al., 2013; Kühner et al., 2007). A tool that has been standardized in many different languages and cultures allows for cross-cultural comparisons (Howitt & Cramer, 2017). However, when adapting or translat- ing questionnaires, it is important to discuss and test the extent to which these modifica- tions may impact the reliability and validity of the questionnaire, as determined by the original version of the tool.

Sampling is an important aspect that receives significant attention in survey research. Sur- veys typically involve large random samples to gather precise and detailed information for the purpose of generalizing findings (Cuttler et al., 2019). Surveys can vary in size (ranging from short to long), approach (face-to-face, telephone, or internet), and have almost no restrictions in terms of the areas of investigation (covering topics from voting intentions to mental health, as long as data is collected through a set of questions) (Cuttler et al., 2019). Data collected in surveys is analyzed both quantitatively (using statistical analysis) and qualitatively (not relying on statistical analysis). The research approach based on surveys is generally considered nonexperimental Regardless of the purpose, surveys remain one of the most important data collection methods in psychological research. They are valuable for collecting data to gain a better understanding of health-related issues, such as estimat- ing mental health during the pandemic in Europe (Wirkner et al., 2021) or worldwide esti- mates on the same topic (World Health Organization, 2022).

##### Survey Development

In a study on anxiety, the selection of a questionnaire depends on the specific focus of the research question. There are questionnaires available that measure fear related to specific situations or objects (phobias), as well as questionnaires designed to assess generalized anxiety disorders. The appropriate questionnaire to choose will depend on the relevance to the research question. It is worth noting that there may be competing questionnaires available, which means they are designed to measure the same construct (e.g., general- ized anxiety disorder).

For example, in 2018, two research groups, one in Australia (Schilpzand et al., 2018) and one in the UK (Williamson et al., 2018), independently developed questionnaires and con- ducted surveys to measure parental unfavorable thinking styles in relation to their child- ren’s traumatic experiences. Each research group took a different approach to developing their questionnaire. One group utilized classic strategies of questionnaire construction (Schilpzand et al., 2018), while the other selected questions based on expert evaluations in the field (Williamson et al., 2018). To select one of these questionnaires, researchers should carefully examine the individual questions and the underlying concept behind them. Additionally, it is important to compare the reliability and validity of the two ques- tionnaires.

**Item** An item refers to a ques- tion or statement formu- lated in a questionnaire.

A survey is a compilation of multiple questionnaires or individual questions. While its administration is relatively straightforward, constructing a survey can be challenging. These challenges primarily revolve around the responses and how they are influenced by various factors, including the wording of the questions, the order in which the **items** are presented, and the available response options. Additionally, the psychological processes involved in responding to questions can also impact the responses. It is important to rec- ognize that responding to questions is not as simple as it may seem. Each participant must read, interpret, make judgments, and then convert their judgments into a response that aligns with the available response options (Cuttler et al., 2019).

For instance, consider the following question taken from a questionnaire designed to assess alcohol use disorder (Babor et al., 2001): “How often during the last year have you been unable to remember what happened the night before because of your drinking?” The response options provided are: 0 – never, 1 – monthly or less, 2 – monthly, 3 – weekly, 4 – daily or almost daily (Babor et al., 2001)*.* While the question may appear simple, partic- ipants must read and comprehend what is being asked. However, the formulation of the question is not straightforward. Participants need to not only focus on the previous night but also consider their drinking pattern over the past year. After understanding the ques- tion, they must retrieve relevant information from their memory. This can be challenging as the question pertains to events that occurred over the past year, requiring participants to estimate and recall instances of memory loss beyond recent drinking occasions. Once they have made a reasonable estimate, they need to select the response option that best reflects their drinking pattern. Additionally, participants may also face the dilemma of pro- viding an honest response, particularly if they believe their drinking is causing memory loss more frequently than they are willing to admit. Researchers must carefully consider and assess these issues when creating a survey to ensure that the findings are not biased or ambiguous (Cuttler et al., 2019).

Each questionnaire consists of a series of questions or statements known as items, which correspond to judgments. When preparing items, several considerations need to be taken into account. The wording of each item should be clear, simple, and easily understanda- ble, avoiding ambiguous terms or jargon. It is important to consider the target group and formulate items appropriately for their level of understanding. It is generally recom- mended to create or use items that do not impose excessive mental or cognitive demand on respondents. When the study aims involve past experiences, it is good practice to spec- ify the exact period of interest, such as the past two weeks or the past 12 months. It is advisable to avoid questions that force a specific answer, such as forced-choice items that limit respondents to selecting one of two choices (typically yes or no). Additionally, items that include double negatives should be avoided, as they can be confusing and require extra time for respondents to interpret the intended positive meaning (Cuttler et al., 2019).

Additionally, the questionnaire instructions must be informative and easily understood. The order in which the items are presented could also impact the respondent’s behavior. It is recommended to place questions that are too difficult or personal towards the end of the survey so that the respondent can become more accustomed to the process and not disengage at the onset. Standardized questionnaires typically offer a rating scale as answer options for items. Rating scales depict sections of a characteristic continuum (Dör- ing & Bortz, 2016) and refer to frequency (e.g., never, rarely, occasionally, often, always), intensity (e.g. not at all, hardly ever, moderately, mostly, completely), probability (e.g. not at all, probably not, maybe, fairly likely, definitely) or evaluation (e.g., does not apply at all, hardly applies, partly applies, quite applies, fully applies).

##### Rating Scales

A rating scale refers to a common question type, typically closed, that includes different answer options with varying weights. In a survey, respondents are typically asked to select from several options that are scaled to represent two extremes or varying degrees (e.g., from totally happy to extremely unhappy) (Shaughnessy et al., 2012). Rating scales pro- vide a subjective way to measure responses related to levels of satisfaction, feelings, or experiences. For instance, the IU wants to assess the overall satisfaction of its students with the university’s services. The generic item created to measure satisfaction could be “On a scale from 1 to 5, how satisfied are you with IU’s services?” The response options provided would hold a certain value assigned to each option (e.g., 1 – extremely dissatis- fied, 2 – dissatisfied, etc.). Rating scales commonly use response options ranging from 1 to 10, 1 to 7, or 1 to 5. The **Likert scale**, which uses response options from 1 to 5, is the most widely used scale in surveys. In a rating scale, the response option 5 represents the positive end, while the response option 1 represents the negative end (and never vice versa).

It is important to distinguish between a rating scale and a psychometric scale. A rating scale refers specifically to the response format used in a questionnaire, while a psycho- metric scale encompasses a group of items (questions) that are answered using the same rating scale. When constructing a questionnaire, researchers often start with a large pool of questions derived from the current state of research. Respondents then complete this lengthy questionnaire. Statistical methods are employed to determine which items are appropriate indicators of the content-related phenomenon being studied. Items that are

**Likert scale**

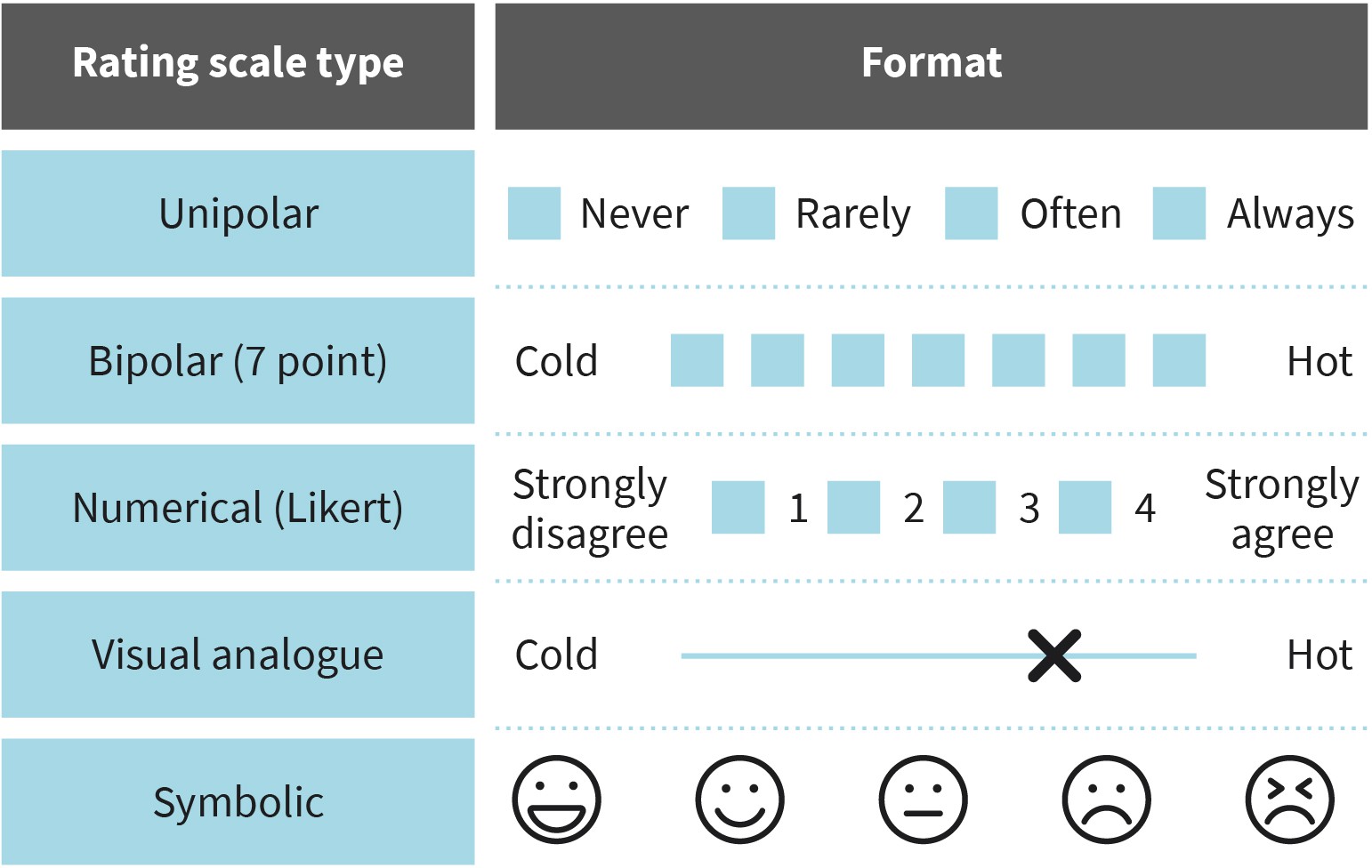
This is a type of rating scale that is commonly used in research with questionnaires.

related to each other are combined into a single factor, representing a measure of the latent construct (Fricker et al., 2012). Factor analysis is a statistical method commonly used to identify latent traits from survey questions (Fricker et al., 2012). Rating scales used to measure psychological phenomena and dimensions are typically considered ordinal scales of measurement. However, it is often assumed that the distances between the response option values are equal, leading to the treatment of ratings as interval scales (Shaughnessy et al., 2012).

In the field of clinical psychology, several questionnaires are evaluated using cut-off val- ues. These values are set to identify individuals with particularly high levels of a specific characteristic. Questionnaires are commonly used as screening tools in clinical psychol- ogy to determine which individuals have clinically abnormal values and require further diagnostic assessment. For example, the Beck Depression Inventory (BDI-II) is frequently used to diagnose depression. If an individual’s scores on the BDI exceed the cut-off score, it typically indicates the need for a clinical diagnostic interview. Cut-off values are estab- lished to strike a balance between sensitivity and specificity in diagnostic judgments. Sen- sitivity refers to the ability of a questionnaire to accurately identify individuals who have depression. Specificity, on the other hand, refers to the ability of a questionnaire to cor- rectly identify individuals who do not have depression among those who have tested neg- ative for depression. In the case of depression, research has indicated that a cut-off value of 14.5 for the BDI-II is considered optimal. This cut-off value has been found to have a sensitivity of 0.86 and a specificity of 0.78 (von Glischinski et al., 2019). A sensitivity of 0.86 means that out of 100 people with depression, 86 are correctly identified as having depression. The specificity value of 0.78 indicates that out of 100 healthy individuals com- pleting the BDI-II, 78 are correctly identified as not having depression (von Glischinski et al., 2019).

When collecting data using rating scale response options, it is common to encounter dis- tortions in the collected data caused by respondents’ response biases (Cuttler et al., 2019; Shaughnessy et al., 2012). Response biases occur when survey participants do not provide honest or accurate responses. There are several types of response biases that can impact the data. One common response bias is demand bias, which arises from participants being influenced by the act of participating in a survey or research study. As a result, their responses may not truly reflect their genuine opinions or beliefs. Social desirability bias occurs when participants provide answers that they believe are socially desirable rather than expressing their true thoughts or feelings. Data distortions can arise from extreme responses, where participants consistently provide extremely positive or negative answers, or when respondents always select the neutral response option. Other biases include acquiescence and dissent, which reflect a tendency to agree with all survey items or disagree, respectively.

Figure 9: Rating Scales



Source: Evangelos Zois (2022).

#### Psychometric Measures Derived from Quantitative Methods

To attain a high level of proficiency in quantitative research methods and to systematically assess the potential drawbacks of the quantitative approach, specific criteria for meth- odological quality have been introduced. These criteria pertain to objectivity, reliability, and validity (Cuttler et al., 2019).

Scientific objectivity refers to the absence of bias in all aspects of a research study, includ- ing research questions, methodology, statistical analysis, interpretation of results, dissem- ination of findings, and the investigators themselves. A high level of objectivity is achieved when different individuals conduct the same study under identical conditions and arrive at the same conclusion. This indicates that the study, tools, and evaluation were all con- ducted objectively, following standardized procedures (Cuttler et al., 2019).

##### Reliability

In psychological research, the concept of reliability pertains to the consistency and stabil- ity of both a research study and the measurement tool used. Reliability is closely linked to replication, as it signifies the ability to reproduce findings consistently (McLeod, 2007). For instance, when the results of a research study are replicated multiple times, their reliabil-

ity can be confidently asserted (Cuttler et al., 2019). There are various types of reliability, including test-retest and interrater reliability (external), as well as internal consistency (internal).

###### Test-retest reliability

It refers to the consistency of results obtained from a repeated survey using the same instrument. For instance, in a study assessing IQ, participants are administered an IQ test on two separate occasions, with a 45-day interval between assessments. The extent of cor- relation between the two IQ scores serves as an indicator of the test-retest reliability of the IQ test.

###### Interrater reliability

The interrater reliability refers to the agreement between assessments conducted by dif- ferent observers or raters. For instance, in a university setting, a group of lecturers is tasked with evaluating the writing skills of master’s students using method A. Subse- quently, a separate group of lecturers from the same university is asked to assess the same students using the same method A. The level of similarity in the evaluations pro- vided by both groups would serve as an indicator of interrater reliability.

###### Internal consistency

Internal consistency refers to the degree of association between the questions of a survey and the corresponding answers, ensuring that they measure the same characteristic and minimize measurement error (Cuttler et al., 2019; McLeod, 2007). For instance, if a survey respondent agrees with the statements “I enjoy swimming” and “I have enjoyed swim- ming in the past,” while expressing disagreement with the statement “I hate swimming,” this would indicate good internal consistency for the questionnaire employed.

##### Validity

It refers to the extent to which a study accurately measures what it was intended to meas- ure, indicating the degree of correspondence between the collected data and the origi- nally intended concept (Cuttler et al., 2019; McLeod, 2013). There are different forms of validity, including content validity, construct validity, and criterion validity.

###### Content validity

Content validity refers to the degree to which a measurement accurately represents the most effective operationalization of a characteristic. For instance, a depression assess- ment tool would have low content validity if it only captures the emotional aspects of depression but fails to account for the physical symptoms associated with the condition.

###### Construct validity

This refers to the extent to which a measurement tool or test accurately measures the the- oretical concept it is intended to assess (McLeod, 2013). In the case of the newly con- structed questionnaire to evaluate depression, the construct validity of the tool would be determined by its ability to specifically measure depression and not other related condi- tions such as dysthymic disorder or post-traumatic stress disorder (PTSD).

###### Criterion validity

Criterion validity refers to the extent to which the results of a measurement instrument align with those of another already validated and widely accepted instrument (Cuttler et al., 2019; Nikolopoulou, 2022). For instance, if a researcher is comparing a newly devel- oped questionnaire designed to assess depression with an already established and well-validated tool that measures the same construct, they are examining the criterion validity of the new questionnaire (Nikolopoulou, 2022).

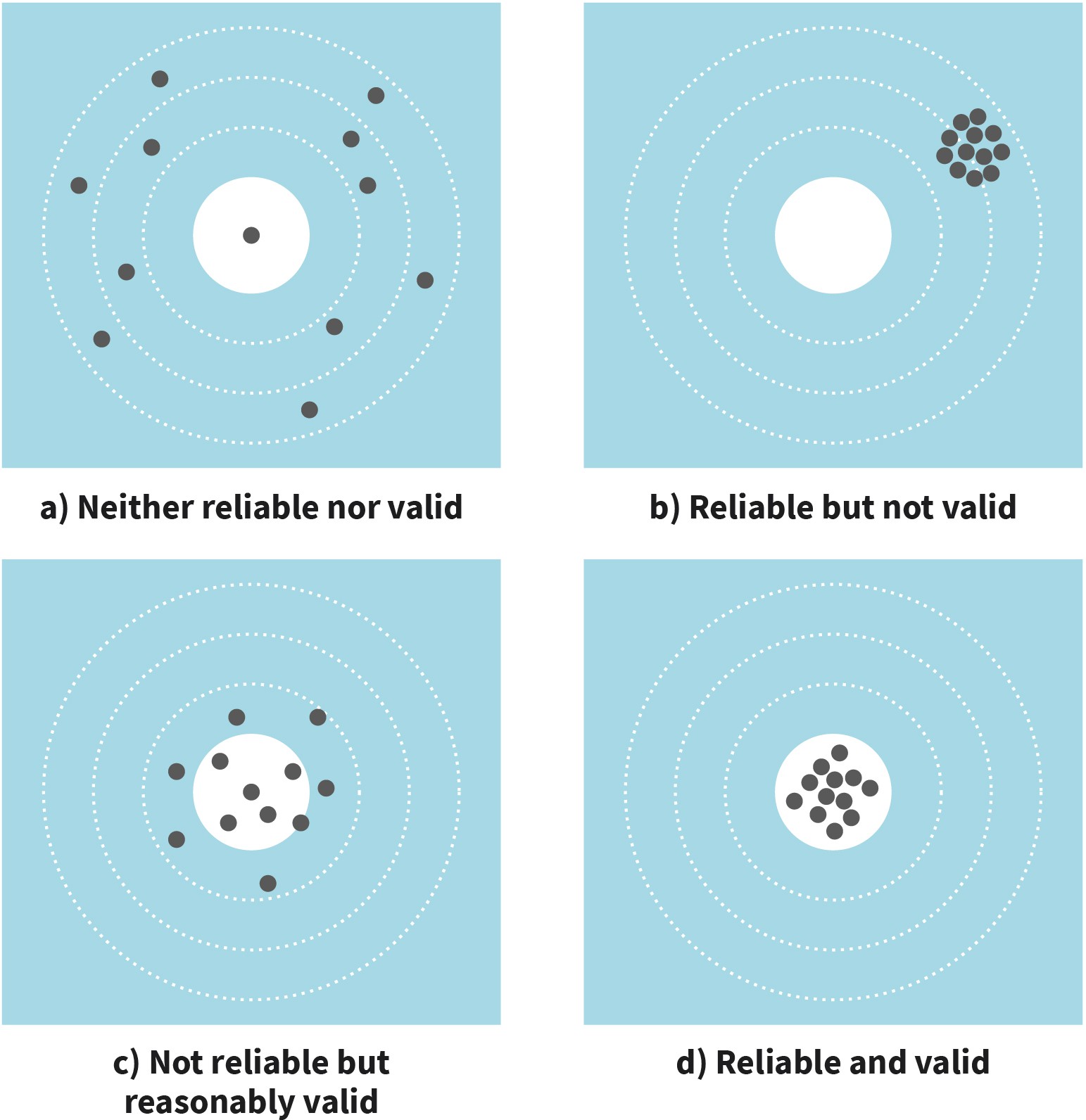
##### Objectivity

Objectivity serves as a fundamental requirement for the other quality criteria. Without objectivity, reliability and validity cannot be achieved. A study can be considered valid if it is both objective and reliable (Hussy et al., 2013). However, it is possible to have highly reliable instruments and still not obtain valid data. For instance, consider a weighing scale that consistently adds 2 kilos to your actual weight. Although the scale is reliable because it consistently adds the same 2 kilos, it is not valid because it introduces an inaccurate measurement by adding the extra weight.

An analogy to illustrate the achievement of reliable and valid results can be seen in a sim- plified version of the game of darts, as depicted in the figure. The objective of this game is to throw darts at a dartboard hanging on a wall, much like the goal of obtaining reliable and valid results. By conducting the test multiple times, the results can be compared, as shown in the figure.

* When darts scatter randomly across the board, the results are neither reliable nor valid (a).
* If the darts consistently land in a similar area but are far from the target, the results can be considered reliable but not valid, as they do not achieve the intended goal (b).
* When the darts cluster around the center of the dartboard, the results may not be relia- ble, but they are reasonably valid (c).
* When all the darts hit the center circle (target) of the dartboard, the results are both reli- able and valid (d).

Figure 10: Reliability and Validity of an Investigation



Source: created on behalf of IU (2023).

#### Psychological Testing

A psychological test is comprised of a series of tasks that have been empirically tested and fully standardized in order to measure individual characteristics (Döring & Bortz, 2016). Individuals with different skills or characteristics can participate in these tasks and their performance may vary. The goal of psychological tests is to objectively measure character- istics that cannot be directly observed or measured, such as intelligence and personality traits (latent variables). These tests can focus on a single aspect of a construct or have multiple dimensions. An example of a multidimensional psychological test is a psychopa- thology inventory, like the Minnesota Multiphasic Personality Inventory (MMPI) (Camara et al., 2000). The MMPI is a comprehensive psychological assessment tool used in clinical and legal settings. Its purpose is to evaluate personality characteristics and identify poten-

tial psychiatric conditions, including abnormalities in personality. In contemporary times, this tool is utilized to evaluate individuals in the aviation and police sectors prior to their initiation of duty. The tool consists of 10 subscales, each of which assesses various aspects of both normal and abnormal behavior. One of the dimensions included in this test is referred to as Masculinity/Femininity. In a broad sense, this dimension measures the extent of masculinity and femininity in individuals, with a focus not solely on sexuality. It assesses individual interests, hobbies, sensitivity, aesthetic preferences, and other related factors (Camara et al., 2000).

In addition to the aforementioned psychological tests, there are other types of assess- ments used in various contexts, e.g. to evaluate specific abilities such as mathematical or perceptual skills (Aiken, 1998). Clinical tests, on the other hand, are utilized in clinical set- tings to aid in diagnosing conditions, such as the previously mentioned MMPI test for per- sonality assessment (Camara et al., 2000), or the Beck Depression Inventory (BDI) for eval- uating symptoms of depression (Beck et al., 2011). Neuropsychological tests focus on measuring cognitive abilities, including memory, decision-making, and perception. These tests are often conducted to provide additional diagnostic information for individuals who have experienced brain injuries, and are typically administered by clinicians such as neu- rologists (Lezak & Lezak, 2004). Personality assessments are another category of psycho- logical tests that evaluate constructs related to personality traits, such as the Big Five per- sonality test. These assessments aim to capture various elements of an individual’s personality (Soto & Jackson, 2013). Furthermore, psychological tests are frequently uti- lized in human resources departments to evaluate an individual’s suitability for a particu- lar position. In research settings, psychological tests are routinely employed to collect inter-individual characteristics in an objective manner. Intelligence tests represent another category of psychological assessments, typically designed with progressively increasing difficulty levels for the items. One commonly administered intelligence test is the Wechsler Intelligence Test (WAIS) (Wechsler, 2012). This test consists of various sub- scales that assess language comprehension, perception, logical thinking, and working memory (Wechsler, 2012).

Throughout this unit, the term “standardized” has frequently occurred. In order for psy- chologists or researchers to utilize any of the aforementioned tests, it is crucial that the tests are standardized (StudySmarter, n.d.). Standardization entails using a test that is reli- able, accurate, and includes clear instructions for administration. Additionally, the test scores should be derived from the average population, allowing psychologists or research- ers to compare their sample scores to the average population test scores. The average test scores obtained from the population and included in a reliable test are referred to as norms or normative data. The process of obtaining norms is a significant aspect of psycho- metrics in psychology (StudySmarter, n.d.). For example, consider a scenario where a psy- chologist is administering an IQ test to a sample of 40 individuals. In order to interpret the scores accurately, the psychologist needs to compare them to the average score of the population that shares the same characteristics as the sample. For instance, if the sample consists of 40 male participants between the ages of 25 and 35, the psychologist would refer to the norm score for that specific age group and gender. The norms provided with the IQ test used in this example would inform the psychologist that the average IQ score

for men between the ages of 25 and 35 is 100. With this information, the psychologist can then categorize the IQ scores of each individual in the sample as either low, average, or high, based on the understanding that a score of 100 represents the average IQ.

Psychological tests vary in terms of their theoretical foundations. The majority of tests in psychology are based on the classical test theory (CTT) (Cappelleri et al., 2014). CTT is a quantitative approach used to assess the validity and reliability of a measurement tool based on its items or questions. According to CTT, each observed score (X) is a combina- tion of an underlying true score (T) of the construct being measured and random error (E) (Cappelleri et al., 2014). Every individual possesses a true score (T) that would be obtained if the measurement tool had no errors. The true score of an individual is consid- ered the expected score over an unlimited number of independent administrations of the measurement tool (Cappelleri et al., 2014). However, when data is collected using a meas- urement tool, the true score of each person is not obtained. Instead, an observed score is obtained, which is accompanied by a certain degree of error (Cappelleri et al., 2014):

observed score X = true score T + error E

#### Biophysiological and Neurophysiological Measurements

One of the significant challenges in psychology is the objective measurement of various characteristics. An approach to address this challenge is by capturing biophysiological and neurophysiological markers of human thought, experience, and behavior. Through the application of relevant methods, modern psychologists are able to detect physiological or biological activity originating from the human body and attribute those signals to specific behaviors, thoughts, or emotions. These methods can examine activity in the central nerv- ous system and the human brain, such as electroencephalogram and neuroimaging tech- niques. Additionally, there are methods that record physiological markers generated from our skin, muscles, sensory organs, cardiovascular system, as well as the hormone and immune systems (Hussy et al., 2013). A summary of selected and commonly used biologi- cal, neurophysiological, and psychophysiological assessment methods will be provided below.

###### Electrodermal activity (EDA)

This phenomenon pertains to changes in the electrical characteristics of the skin that are relevant to the human body. It is also known as skin conductance response (SCR). EDA specifically relates to the electrical properties of the skin, particularly the sweat glands located on the palms of the hands and soles of the feet. When there is an increase in activ- ity in the sweat glands, which can occur due to stress, arousal, or emotional excitement, associated changes in SCR can be measured using appropriate equipment. This typically involves attaching electrodes to the index and middle fingers (Boucsein, 2012). Under- standing these changes in skin conductance can provide valuable insights into the physio- logical responses of the body, helping researchers to better comprehend the complex interplay between emotions and the autonomic nervous system.

###### Eye-tracking

Eye-tracking is a method commonly employed in psychological research to capture and analyze eye movements, eye fixation, and pupil dilation, providing insights into how indi- viduals process visual stimuli. This method utilizes specialized glasses equipped with a high-resolution eye camera. Eye movements are considered indicators of perceptual proc- esses, such as object recognition and differentiation, while pupil dilation is seen as an indicator of arousal (Mele & Federici, 2012).

###### Electrocardiography (ECG)

Electrocardiography (ECG) is a technique used to measure changes in the electrical activ- ity of the heart muscle, specifically the action potentials, which is reflected in the heart rate. This method involves attaching electrodes to the chest area to capture the electrical signals produced by the heart. Heart rate serves as an indicator of stress and emotions.

###### Measurement of the endocrine system

The endocrine system can be measured by assessing hormone levels in various bodily flu- ids, including blood, saliva, urine, or hair. These biological measurements allow for the determination of hormone levels and can provide indications of various physiological processes, including stress. Changes in hormone levels can serve as indicators of stress, among other factors.

###### Electroencephalogram (EEG)

Electroencephalography (EEG) is a neurophysiological technique used to record and measure the electrical activity of the brain. This process involves placing metal disc elec- trodes on the surface of the skull to capture the electrical potentials generated by the brain. EEG signals provide valuable information about brain activity, including processes related to perception, attention, and emotion. It is important to note that electrical brain activity is present even during sleep, making EEG a useful tool for studying various aspects of brain function.

###### Functional Magnetic Resonance Imaging (fMRI)

Functional magnetic resonance imaging (fMRI) is a noninvasive brain imaging technique that has provided scientists with a detailed view of the human brain in vivo (Jezzard, 2008; Logothetis & Wandell, 2004). fMRI relies on changes in blood oxygen levels as an indirect measure of brain activity. When a specific area of the brain is engaged in a task, there is an increase in blood flow to that region (Jezzard, 2008; Logothetis & Wandell, 2004). This increased blood flow is associated with changes in oxygen levels in the same region, which in turn affects the magnetic resonance (MR) signal. The MRI machine can detect and record these changes, providing an indirect depiction of active groups of neurons (Jez- zard, 2008; Logothetis & Wandell, 2004).

###### SUMMARY



In quantitative research, the primary objective is to collect human char- acteristics in a precise and objective manner, and then analyze them quantitatively. To achieve this, reliable and valid instruments are utilized to accurately gather information on these characteristics. These instru- ments can include questionnaires, psychological tests, as well as physio- logical and neuropsychological measurement methods.

Questionnaires typically consist of multiple items that are answered using a rating scale with multiple levels. The sum of the item values on the scale, such as a Likert scale, is then used as a measure of the corre- sponding characteristic in the analysis.

Psychological tests are employed to assess characteristics that cannot be directly observed or inquired about. These tests involve a series of standardized tasks, and individuals with different characteristics may perform these tasks with varying degrees of success. This allows for the determination of an individual’s value and facilitates comparisons with others. Psychological tests often measure cognitive abilities, such as emotions, memory, and intelligence.

In addition to psychological tests, physiological non-neuropsychological tools can be used as markers for psychological variables, such as stress and excitement. These measurements can directly capture the activity of the central nervous system through imaging methods, or the activity of organ and hormone systems outside of the central nervous system, such as skin conductance.

# UNIT 5

## QUALITATIVE AND MIXED RESEARCH

**METHODS**

**STUDY GOALS**

On completion of this unit, you will be able to ...

* name the principles of qualitative research methods.
* describe data collection methods in qualitative research.
* define grounded theory and other qualitative analysis methods.
* explain the criteria of qualitative research.
* outline how qualitative and quantitative methods can be combined.

### 5. QUALITATIVE AND MIXED RESEARCH

**METHODS**

#### Introduction

A relatively recent research study was conducted in the Netherlands, focusing on young individuals who had experienced a traumatic event, specifically one parent murdering the other (Alisic et al., 2017). Such distressing events are associated with significant suffering for those affected, particularly the children, and pose significant challenges for the institu- tions responsible for providing care. The study employed a qualitative approach, with the research team conducting semi-structured interviews with 23 affected children and utiliz- ing thematic analysis to derive their findings (Alisic et al., 2017). At the outset of the project, there were only limited studies available on the mental health of affected chil- dren, with previous research primarily focusing on legal aspects related to custody deci- sions. The emotional conflicts experienced by the affected children can be highly intricate, as they grapple with the loss of one parent while also coming to terms with the fact that the surviving parent committed the murder (Alisic et al., 2017). The researchers opted for the qualitative research method approach due to the nature of the topic and research question at hand. Over a span of ten years, from 2003 to the end of 2012, approximately 140 cases of such nature were documented in the Netherlands, indicating a relatively small population of interest. Qualitative research methods are especially well-suited for research questions and topics that have limited existing empirical knowledge, as well as for complex subjects that delve into the human experience.

#### Principles of Qualitative Research

Qualitative research differs significantly from quantitative research in its underlying assumptions, methods, and procedures. As a result, there are distinct characteristics or principles that typically apply exclusively to qualitative research methods (Hancock, 1998; Howitt & Cramer, 2017). However, before delving into the description of these principles, it is important to provide a definition of qualitative research. Qualitative research involves the collection and analysis of non-numerical data, such as text or audio, with the aim of gaining in-depth insights into people’s opinions and experiences, or generating new research ideas and concepts (Bhandari, 2022). Unlike quantitative research, qualitative research is not primarily associated with numerical data and statistical analyses (Bhan- dari, 2022; Howitt & Cramer, 2017).

The following principles or characteristics differentiate qualitative research methods from quantitative research methods (Hancock, 1998; Howitt & Cramer, 2017): Qualitative research

* + - investigates the opinions, experiences, and feelings of individuals, generating subjec- tive data.
* describes events and phenomena as they occur in a natural setting (no lab). There is no manipulation of the condition under investigation as in the case of quantitative research.
* aims to comprehend a situation by adopting a holistic approach which is not bound to the identification of a set of variables.
* follows an inductive approach, utilizing data to develop new concepts or theories and gain a deeper understanding of the natural world.
* employs data collection methods that involve direct interactions with individuals, such as conducting interviews with individuals or groups, as well as engaging in observation.
* involves collecting data in a meticulous and time-consuming manner. While the volume of data collected in qualitative research can be substantial, the sample sizes are often small.
* utilizes sampling techniques to gather data from specific groups and subgroups within the population. In qualitative research, the emphasis is not on having a random or rep- resentative sample of the general population, making it less of a concern (Hancock, 1998).

The following table offers a comparison between qualitative and quantitative research.

Table 4: Qualitative and Quantitative Research Methods Comparison

**Qualitative Quantitative**

Source: Evangelos Zois (2023), based on Hancock (1998).

|  |  |
| --- | --- |
| Subjective | Objective |
| Holistic | Reductionist |
| Phenomenology | Scientific |
| Descriptive | Experimental |
| Naturalistic | Generated |
| Inductive | Deductive |
| Anti-positivism | **Positivism** |

In addition to the aforementioned fundamental concepts, the following principles of qual- itative research are provided below, highlighting the distinctions between qualitative and quantitative research method approaches (Howitt & Cramer, 2017). These principles serve as a guide for researchers to delve deeper into the subjective experiences, meanings, and interpretations of individuals, allowing for a more nuanced understanding of complex phenomena that may not be easily quantifiable.

**Positivism**

In social sciences, positi- vism refers to the per- spective that advocates for the use of experi- ments, statistical meth- ods, and analyses in

research investigations.

###### Response options

In qualitative research, open methods, such as open-ended questions in an interview, are predominantly employed without predefining categories. In contrast, quantitative research utilizes fixed categories for responses (e.g., on a rating scale like Likert) or codes data using predefined categories (true or false). In qualitative surveys, participants have the freedom to provide their answers in their own words.

###### Sample size

In qualitative research, a limited number of cases are examined in depth-. Changes over time are often captured through one-time surveys to create a comprehensive overview. In contrast, quantitative research focuses on examining selected characteristics within a larger sample, typically with a time limited- expression.

###### Inductive approach

Qualitative research frequently adopts an inductive approach, generating theories and hypotheses from an unexplored subject area with limited existing information. In contrast, quantitative research typically follows a deductive path, seeking to confirm or refute exist- ing knowledge.

###### Design flexibility

Due to the inductive approach, qualitative research often uncovers new insights during the survey process. These new findings may lead to decisions to modify the research pro- cedure or instruments. Such changes are not commonly made in deductive quantitative research and are viewed as potential risks to objectivity.

###### Data interpretation

Qualitative research primarily aims to describe and understand phenomena in depth, while quantitative research focuses on establishing causal relationships between varia- bles.

###### Researchers as a measuring tool

Qualitative research often involves a high level of interaction, with researchers serving as direct data collection tools during interviews or observations. The data collected in quali- tative research heavily relies on the researcher, including their understanding of the topic being investigated, their opinions, and their expectations. This characteristic of qualitative research acknowledges that thoughts and interactions play a crucial role in the interpreta- tion of the results.

###### Generalization

Qualitative methods are not primarily focused on generalizing findings to the population; instead, their aim is to achieve analytical generalization of results to theory. In contrast, quantitative research aims to generalize results from a sample to the population in order to confirm theories or hypotheses.

Qualitative and quantitative research methods are not mutually exclusive alternatives in scientific psychology; rather, they are complementary methods suitable for studying dif- ferent objectives. Throughout history, there have been phases where qualitative methods have gained significance and application. However, most psychological departments in Europe and the US primarily conduct quantitative research.

#### Qualitative Methods

The research design in qualitative research plays a crucial role in defining the research goal and process. It encompasses the framework conditions for contact and communica- tion between researchers and participants. Given the diverse nature of qualitative research, study designs vary depending on the research question and the available options for accessing information. One of the most common research designs involves conducting individual qualitative interviews with a selected, relatively small sample. In this section, the major research designs commonly employed in qualitative research will be described.

##### Qualitative Research Designs

The focus of this section is on grounded theory, ethnography, action, field, phenomeno- logical research, case study, and document analysis.

###### Grounded theory

Grounded theory is characterized by its primary focus on developing a new theory about a specific phenomenon through the collection and analysis of data (Howitt & Cramer, 2017). This qualitative research design aims to gather information that will contribute to the development of this new theory, offering fresh insights into known and existing problems (Howitt & Cramer, 2017). For instance, grounded theory has been instrumental in helping clinicians understand the process of grief. Through observations, researchers have gained knowledge about the five stages of grief and the specific reactions associated with each stage (denial, anger, acceptance, resolution). While grief is not a new experience, research utilizing grounded theory has contributed to acknowledging and describing this phenom- enon. Furthermore, the new knowledge derived from grounded theory has enhanced our understanding of grief and how it is experienced by different individuals (Hancock, 1998).

###### Ethnography

Ethnography, originally rooted in anthropology, is a research method commonly employed when a researcher aims to gather data on individuals belonging to a specific culture. The concept of culture is central to ethnography, as it implies that the group of

people being studied share certain characteristics, such as geographical location or reli- gious experiences (Hancock, 1998; Howitt & Cramer, 2017). For instance, in a mental health setting, the ethnographic approach might be chosen when therapists suspect that the cultural background of certain patients is influencing their response to treatment. Research has demonstrated that culture plays a crucial role in shaping the understanding of mental illness, symptom recognition, and response to treatment (Kleinman, 1988). For example, individuals in Asian cultures tend to focus on bodily symptoms rather than the emotional aspects of mental health conditions, which can potentially impact their response to treatment (Lin & Cheung, 1999).

###### Action research

Action research involves collaborative efforts between researchers and those affected to investigate potential solutions to significant social problems. The fundamental idea is to go beyond presenting research solely in scientific texts and instead transform theory into practice to bring about change (Bhandari, 2022). The process of action research begins with defining the social problem and establishing goals. It then involves iterative phases of data collection and discussions with those affected. Qualitative action research is particu- larly useful for addressing existing social and practical issues. However, due to its subjec- tive nature, action research is not well-suited for testing scientific hypotheses and may not fully meet the requirements of scientific knowledge acquisition. In contemporary practice, action research is often employed within the framework of participatory research, aiming to involve participants as active members rather than mere objects of investigation.

###### Phenomenological research

Phenomenology involves the examination, description, and interpretation of phenomena, such as events, situations, experiences, or concepts, as they are experienced by partici- pants (Bhandari, 2022; Hancock, 1998). It is a valuable approach for understanding exist- ing phenomena that may not have been thoroughly described or examined before.

For instance, while back pain is a well-known phenomenon that has been extensively investigated, research studies have primarily focused on factors such as the types of inju- ries that can cause back pain or the effects of different analgesic drugs. However, phenom- enology can provide insights into the lived experiences of individuals with back pain, including the impact on their family dynamics and how children are affected by having a disabled parent. By exploring the subjective experiences and perspectives of individuals, phenomenology offers a deeper understanding of the phenomenon under investigation (Hancock, 1998).

###### Case study

Single case analysis involves conducting a comprehensive analysis of a pre-defined case, which could be an individual, an event, or any other defined entity. This approach may involve studying extreme cases, typical cases, or cases that hold particular significance. Various materials, such as interviews, files, correspondence, and autobiographies, are included to capture the complexity of the case holistically. As a scientific procedure, indi- vidual case analysis involves several steps, including formulating the research question,

defining the case, selecting relevant materials, analyzing the data, and comparing the case to others for classification purposes. Individual case analyses serve two main purposes: testing methods and conducting in-depth investigations of phenomena that are challeng- ing to comprehend in their entirety and complexity (Howitt & Cramer, 2017).

###### Documentation analysis

In qualitative research, documents such as texts, film or audio recordings, and objects pro- vide clues to human behavior, thought, and experience. After defining the research ques- tion and selecting the relevant documents, a critical assessment of the sources is con- ducted, considering their quality, condition, information content, and relevance to the question. The interpretation of information from the documents follows this evaluation. Document analysis can be combined with other qualitative methods, serving as an addi- tional source of information, particularly when other means of data collection are limited.

###### Field research

Field research is a commonly used qualitative research approach that aims to investigate phenomena in their natural environment. A crucial aspect of field research is establishing trust during the field contact phase. Ideally, the researcher becomes a member of the nat- ural environment and experiences the everyday situations firsthand. The researcher actively participates while maintaining a critical distance. The wealth of information col- lected is analyzed using qualitative-interpretative techniques. Field research is suitable when accessing the natural field does not cause significant disruptions, and ethical con- siderations ensure privacy is respected (Howitt & Cramer, 2017).

##### Data Collection Methods

Data collection methods are carefully chosen to align with the study plan. This subsection highlights the most common methods that involve active participation from the individu- als being studied. These methods include interviews, focus groups, and the collection of naturalistic data through observation. Other methods, such as media analysis and visual data analysis, are also utilized. However, these methods are not described in detail here due to their varying application conditions and levels of systematization, which can be modified based on the research questions being investigated.

###### Interview

The interview is a widely used method for collecting data in qualitative research (Bhan- dari, 2022; Hancock, 1998; Cuttler et al., 2019). Interviews can be structured or unstruc- tured and are typically conducted in person. However, depending on the situation, inter- views can also be conducted over the phone (Cuttler et al., 2019). In structured interviews, an interview guide or script is used, and all participants are asked the same set of ques- tions. The researcher or interviewer rarely deviates from the original script (Cuttler et al., 2019). It is uncommon for the order and wording of each question to vary in interviews. Unstructured interviews typically begin with a small number of general questions, fol- lowed by open-ended questions that allow participants to freely discuss topics of interest. The only part of the interview that may have some structure is the introduction (Cuttler

et al., 2019). The majority of interviews conducted for qualitative research purposes fall into the semi-structured format, which is a combination of structured and unstructured approaches. In this format, the researcher prepares specific questions to ask, but also allows for flexibility to ask more detailed questions based on the topics that arise during the interview, even if they were not part of the original script. The goal is to encourage participants to provide as much elaboration as possible (Hancock, 1998). One limitation of semi-structured interviews is the lack of specificity regarding the length of the procedure and the amount of information collected. Additionally, the sample size in semi-structured interviews is typically small. To facilitate data analysis, the content of the interviews is often recorded and transcribed for later use (Cuttler et al., 2019).

Table 5: Outline of Structured and Unstructured Interviews

**Structured Interview Unstructured Interview**

|  |  |
| --- | --- |
| Analysis of a problem and important elements | Definition of the subject to be discussed |
| ↓ | ↓ |
| Preparation of the interview guide | Elaborating how the narrative will be stimulated |
| ↓ | ↓ |
| Conduct of the interview | Conduct of the interview |
| ↓ | ↓ |
| Recording | Inquiries in the direction of the intended subject structure |

Source: Evangelos Zois (2023), based on Mayring (2016).

###### Focus groups

In qualitative research, collecting data from groups of people can sometimes be more beneficial than from individuals alone. Focus groups are a valuable approach for gather- ing specific information, involving small groups of participants who are interviewed on a particular topic (Hancock, 1998; Howitt & Cramer, 2017; Cuttler et al., 2019). The researcher relies on the interaction among participants in focus groups, as this interaction often reveals more information about the discussed topic compared to one-on-one inter- views. Similar to face-to-face interviews, the content of focus groups is recorded and tran- scribed for analysis purposes (Cuttler et al., 2019). It is important for researchers using this method to be aware of group dynamics that can influence focus groups and their out- comes. This includes considering personality characteristics and individuals who may dominate discussions or participants’ desires to be liked by their peers (Cuttler et al., 2019). According to Hancock (1998), focus groups possess certain characteristics. It is rec- ommended to have a group size of approximately 6 to 10 participants. Having fewer than six people can impede the sharing and collection of information, while having more than 10 participants may make it difficult for everyone to actively participate. Another impor- tant aspect is the number of focus groups that should be conducted within a research project. Hancock (1998) suggests that multiple focus groups should be included to avoid

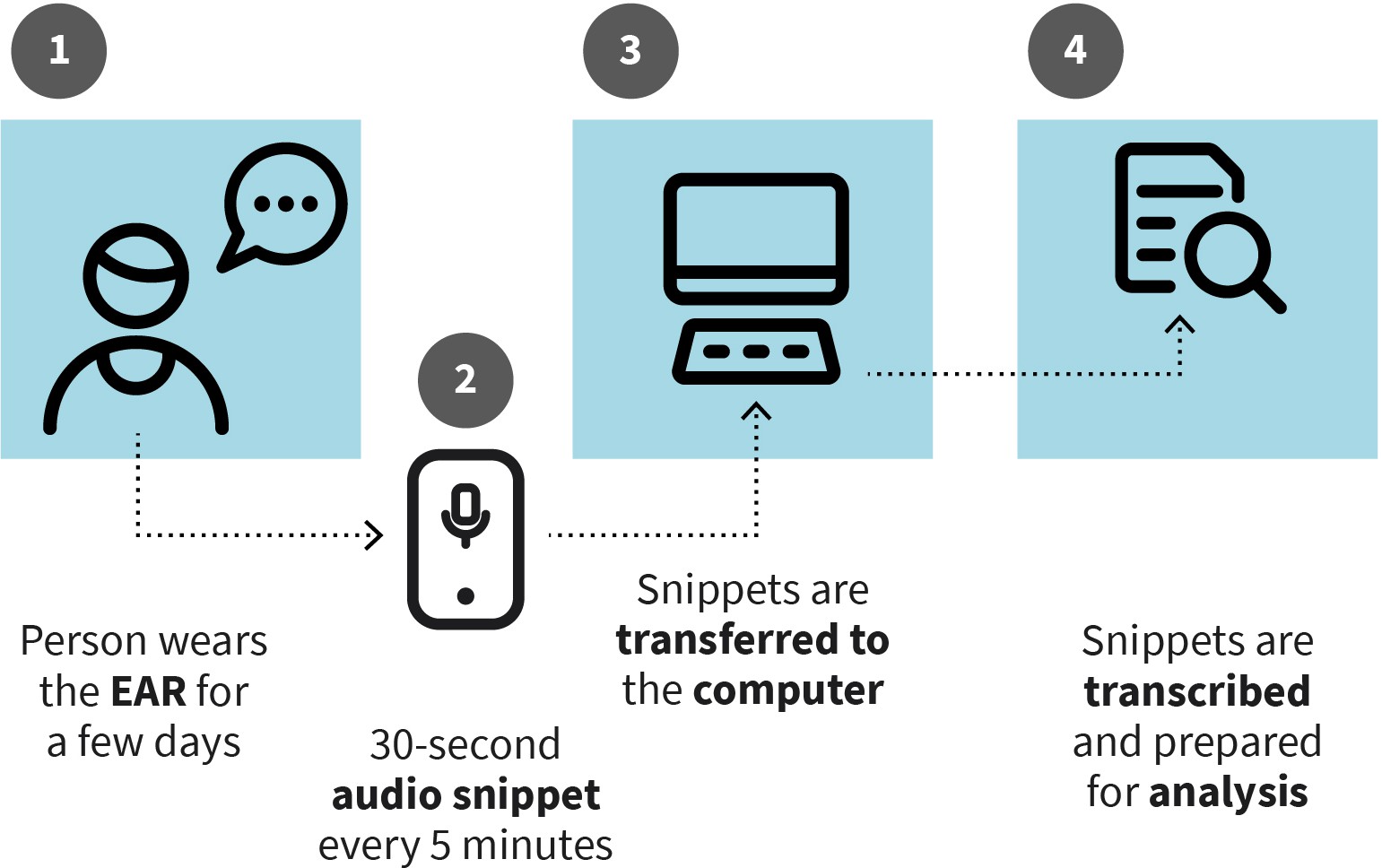
relying solely on one group for data collection. This is because factors within a single focus group can potentially hinder the group’s ability to provide reliable results, such as subop- timal participant interaction that the researcher may not be aware of. In terms of partici- pant selection, it is important for the research to ensure that the participants share com- mon characteristics, such as the same profession or experience (Hancock, 1998).

###### Naturalistic observation

When observation is utilized as a qualitative research design, data is collected in everyday situations, such as conversations, interactions, and behavior, without the active involve- ment of the researcher. This data collection is typically recorded, often through video recording (Hancock, 1998). An example of data collection in a naturalistic setting is the Electronically Activated Recorder (EAR) developed by Mehl et al. (2012). In this method, each participant in the observation is provided with an electronic device, such as an iPod, which is attached to a belt. Participants wear the device for a few days, depending on the study’s requirements. The EAR captures 30-second audio snippets at regular intervals, typ- ically every five to 12 minutes, documenting everyday conversations throughout the day. Research has shown that individuals typically begin discussing the recording device and the study approximately two to three hours after activating the EAR. After this initial period, participants resume their normal daily activities and are not influenced by the knowledge that their conversations are being recorded (Mehl et al., 2012). The collected data is then transferred to a computer and transcribed, allowing for qualitative or quanti- tative analysis.

Naturalistic data obtained through observation can offer valuable insights into everyday social processes. Conducting a naturalistic observation often provides more valuable information compared to other methods, such as interviews. In interviews, there is a pos- sibility that participants may not provide truthful answers, even when asked about their behavior. However, direct observation allows researchers to witness participants’ actual behavior, providing more reliable data collection (Hancock, 1998).

Figure 11: Data Collection Process with EAR



Source: Evangelos Zois (2023), based on Alisic et al. (2017).

#### Qualitative Analysis

Qualitative data can manifest in various forms, including text, photographic material, video, and audio. Therefore, depending on the nature of the research design, a qualitative researcher may work with data derived from interview transcripts (including focus group data), survey responses, or recordings from naturalistic observation (Bhandari, 2022).

##### Handling of Qualitative Data

Irrespective of the origin of qualitative data, there are several useful steps that a researcher can take when preparing data for analysis (Bhandari, 2022; Hancock, 1998). It is essential to organize and prepare the collected data by transcribing (e.g., interviews) or typing notes from observations. Reviewing the data to identify any recurring information or patterns is also crucial. Additionally, creating an appropriate coding system is impor- tant. Based on the initial research ideas, the researcher needs to develop codes that will enable them to categorize the data. The prepared coding system should then be applied to the data, such as assigning designated tags that include the created codes when reviewing responses from each participant. Finally, it is important to characterize the data based on key recurring themes at the end of this process (Bhandari, 2022).

Prior to commencing the actual analysis, it is necessary to transcribe the data into a suita- ble written form. Qualitative data is often in the form of audio recordings, which can be converted into text using appropriate software programs. Nowadays, there are various

technical tools available for recording and transcribing qualitative data. For instance, there are pens designed for taking notes during interviews or focus groups that simultane- ously save the audio document and the notes as a PDF document, with the notes marked with the exact time the audio document was recorded. This allows the interviewer to pre- cisely identify which section of the interview the notes refer to.

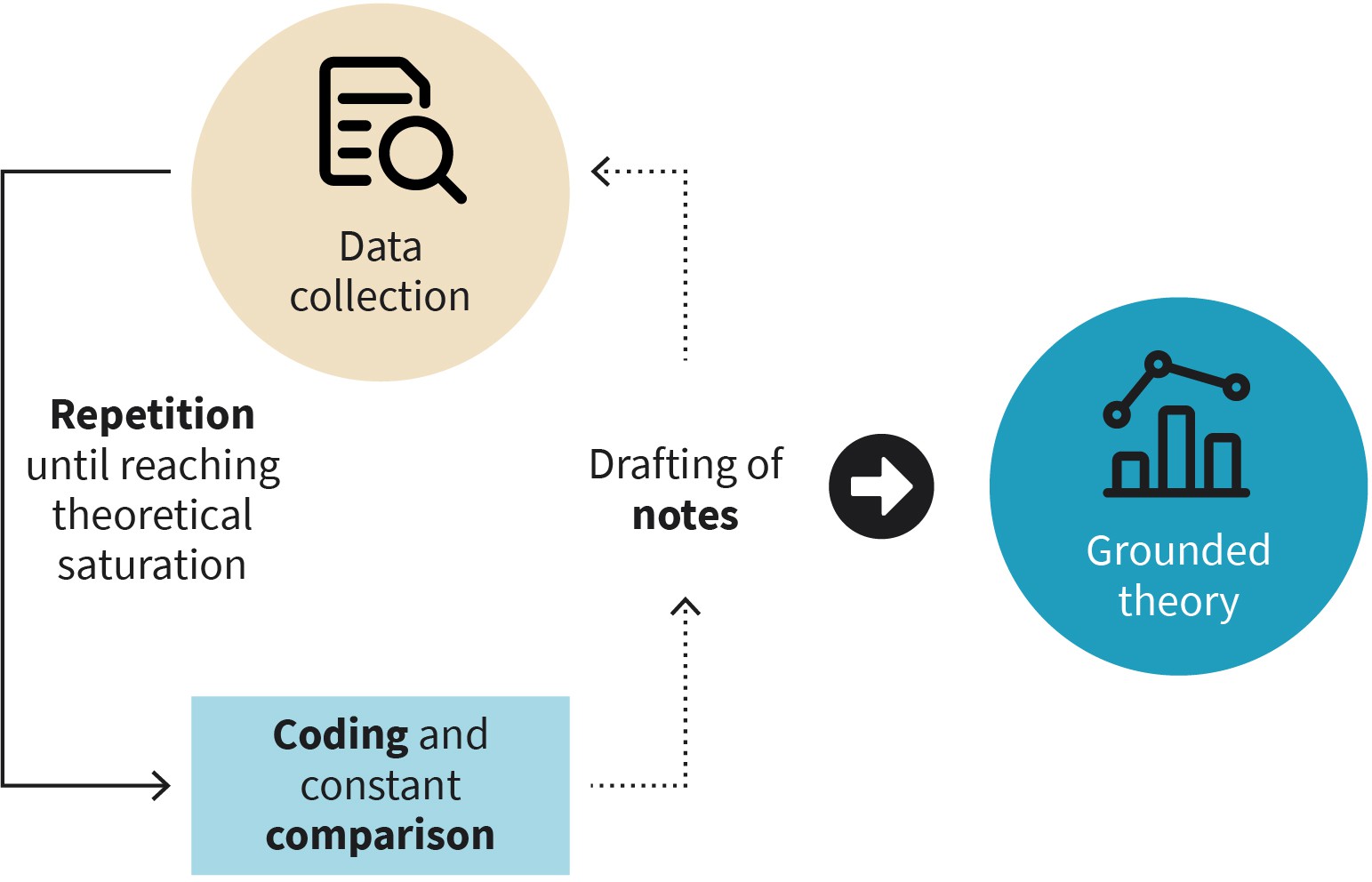
During the transcription process, important decisions need to be made. The recordings can be transcribed holistically or selectively, focusing only on aspects relevant to the research question. Statements can be saved in written, literary transcription, or phonetic transcription. Additionally, it is important to decide whether to transcribe only the text with content or also the co-text, which provides clues on how something is said. The utter- ance may include paraverbal elements (pauses, changes in voice) and nonverbal elements (laughter, yawns). Paraverbal elements are often marked with symbols defined in tran- scription systems (Hussy et al., 2013). Depending on how the data is transcribed, it can be used for qualitative, quantitative, or mixed method- analyses. This allows for the quantita- tive recording of specific words (e.g., motion-related words), the frequency of conversa- tion topics, or the distribution of paraverbal or non-verbal elements to identify connec- tions and patterns. Therefore, it is an important methodological decision to determine which methods are best suited for investigating the research questions.

##### Grounded Theory

In grounded theory, hypotheses, theories, or concepts are developed and modified during the data collection process. Data collection and analysis occur simultaneously. Whenever a theoretically relevant aspect arises during data collection, the researcher creates a note containing their reflection on the collected data and the theoretical concept. These notes can be used to modify the theory or prompt additional data collection. There is a continu- ous comparison between newly collected data and already developed theoretical con- cepts. By collecting new data, researchers assess how well the developed theory can describe the new data. To ensure that the theory is not only applicable to the given sam- ple but also capable of describing or predicting new data, data recollection aims to uncover unexpected data. This approach aims to develop a robust theory that is valid in various contexts (Muller et al., 2016). The process continues until the theoretical concept has been fully formulated and is no longer subject to change by new data, at which point **theoretical saturation** is achieved, and data collection can be concluded (Mayring, 2016).

**Theoretical saturation** This is a term used in grounded theory. It is rel- evant to the stage in data collection when no fur- ther insights or new infor- mation can be derived from the data.

Figure 12: Systematic Process of Theory Formation



Source: Evangelos Zois (2023), based on Forrester (2010).

##### Analysis of Qualitative Data

Analyzing qualitative data involves systematically examining non-numerical information to derive meaningful findings. This method allows for a deeper understanding of complex psychological phenomena.

###### Content analysis

Content analysis is a method used to categorize, classify, and interpret content data. The primary objective of this method is to identify the occurrence of specific words, themes, or ideas within qualitative data, such as text. This data can be derived from various sources, including interviews, focus groups, articles, headlines, or archival documents. In practice, a study utilizing content analysis would involve organizing different texts containing the collected data, and then coding, separating, and classifying the text data into several cate- gories. Once the initial coding process is completed, a second coding phase begins, during which codes are further classified into categories to facilitate additional data review (Han- cock, 1998). An example of content analysis was published by a research group from the University of Bangor, UK (Turnbull et al., 2006). The research group analyzed transcripts from a series of psychodynamic sessions involving a profoundly amnesic patient. The pro- cedure involved reviewing the transcribed sessions the patient had with their therapist. The research team then coded the material, identifying recurring topics in each session, such as everyday life, family, relationships, and familiarity. The objective was to determine whether the patient, despite their amnesia, would demonstrate any indication of remem- bering or recognizing their therapist during each session. Information related to familiarity was coded as a recurring category in almost all sessions. Researchers looked for phrases

or words that would indicate familiarity, such as “I feel comfortable talking to you,” “I don’t feel anxious,” or “I feel very much at ease at the moment” (Turnbull et al., 2006, pp. 201–202).

###### Thematic analysis

Thematic analysis is a method used to condense qualitative data by identifying common and distinct themes. These themes are ideas and concepts that are pertinent to the research question and can appear multiple times in different individuals or even within the same person. Once the themes are identified, they can be analyzed using various theo- retical assumptions. Thematic analysis is considered a fundamental qualitative analysis method and does not require extensive methodological knowledge. Its flexibility and inde- pendence from a specific theory make it applicable in diverse contexts, using various types of qualitative materials such as transcribed interviews, focus groups, naturalistic data, and documents (Bhandari, 2022). For instance, in a focus group discussing the chal- lenges of modern society, 12 participants were involved. The researchers recorded the audio data, transcribed it, and then proceeded to identify themes relevant to climate change and illegal migration.

#### Quality Criteria to Qualitative Research

In this subunit, limitations that are commonly associated with qualitative data and research methodology will be discussed. It is crucial to consider these potential limita- tions for every method or approach in order to make informed decisions about their implementation. Additionally, understanding these limitations allows us to identify meas- ures that can be taken to minimize any disadvantages that may arise.

##### Limitations of Qualitative research

Qualitative research offers a certain level of flexibility, as data collection can occur in natu- ral settings. It has the potential to contribute to the development of new theories and hypotheses, as well as provide fresh insights into existing ideas and concepts (Bhandari, 2022). However, there are limitations associated with collecting, analyzing, and interpret- ing qualitative data. Qualitative data is often considered unreliable for valid reasons. While collecting data through naturalistic observation can be valuable, researchers face the challenge of not being able to control factors that may influence participants’ behavior. This limitation can impact the quality of the data and subsequently affect the interpreta- tion of that data, potentially leading to incorrect conclusions (Bhandari, 2022). Another issue in qualitative research is the challenge of research replication. The researcher’s sub- jectivity and role play a significant role in this matter. The analysis and interpretation of the available data are ultimately determined by the researcher. They decide which aspects of the collected data are important and relevant. Consequently, if the same data were pre- sented to a different researcher, the interpretation and focus of analysis could differ (Bhandari, 2022). Furthermore, qualitative research faces challenges regarding the gener-

alizability of findings. In most cases, the sample size in qualitative research is small, yet it generates a substantial amount of data. Regardless of the volume of data collected, it is extremely difficult to generalize the findings to the broader population. This limitation arises because the samples selected are not random or representative (Bhandari, 2022). Qualitative research has the disadvantage of being time-consuming due to the extensive tasks of handling, transcribing, and analyzing the data. Additionally, researchers involved in this type of research may lack the necessary experience or knowledge. It is crucial for researchers to be well-trained to accurately identify and code the desired concepts (Bhan- dari, 2022).

Due to the differences in objectives, assumptions, and approaches between qualitative and quantitative research methods, it is not appropriate to simply transfer quantitative quality criteria to qualitative research. Attempts to do so have been criticized (Döring & Bortz, 2016), and researchers often face difficulties in establishing universally accepted standards for assessing the quality of qualitative research. In response to this challenge, there have been efforts to formulate specific criteria for qualitative research, with recom- mendations to apply various measures to ensure high quality at each step of the study and to justify the chosen criteria (Döring & Bortz, 2016).

The most used international quality criteria for qualitative research are the four criteria of credibility, which include trustworthiness, transferability, reliability, and confirmability (Lincoln & Guba, 1985). To establish the meaningfulness of qualitative study results, it is important to meet these criteria. Trustworthy results can be transferred to other contexts, are based on a clear research process (reliability) and are free from the researchers’ biases and interests (subjectivity) (Lincoln & Guba, 1985).

**Triangulation** This refers to the process of combining qualitative and quantitative data to improve the validity of research conclusions.

#### Mixed Methods

Qualitative and quantitative methods offer distinct advantages and disadvantages. Rather than directly comparing these two approaches, there has been a growing trend in recent decades to integrate elements of both qualitative and quantitative research. This integra- tion, known as mixed methods research, combines qualitative and quantitative methods to comprehensively investigate a specific research question. The underlying concept behind this mixed methods approach is that it can provide researchers with a more accu- rate and comprehensive understanding of the topic under investigation by leveraging the strengths of both qualitative and quantitative methods (Tegan, 2021). Mixed methods research is appropriate when the researcher believes that neither quantitative nor qualita- tive methods alone are sufficient to explore the research question, often due to concerns about generalizability and the need for a larger sample size to protect external validity. Mixing methods enhances research conclusions by providing contextualized findings and increasing credibility through data collection from the same participants using different methods, ultimately improving validity through **triangulation** (Tegan, 2021).

##### Mixed Methods Designs

There are different kinds of mixed methods designs, and here we will discuss just three of them. The first is convergent parallel design, where qualitative and quantitative data are collected simultaneously but analyzed separately. The overall conclusion is reached by comparing the results from both analyses (Tegan, 2021). The second type is embedded design, where data is collected and analyzed simultaneously, but one approach or data set is supplementary to the other. In this design, the qualitative approach is secondary to the quantitative approach, included to enhance the conclusions drawn from the quantitative analysis (Tegan, 2021). The last mixed methods design discussed here is the exploratory sequential design. In this design, quantitative data collection and analysis are conducted before qualitative data collection and analysis (Tegan, 2021).



**SUMMARY**

Qualitative methods are well-suited for inductively exploring issues with limited empirical knowledge, involving a smaller number of participants to examine their experiences in depth. The interpretations of both respondents and researchers are crucial in data collection and analysis. Qualitative data can be gathered through structured, semi-structured, or unstructured interviews, focus groups, naturalistic observations, and document analysis. Transcribing data accurately is an important step in qualitative research. Common qualitative analysis methods include the- matic analysis, which summarizes and interprets important topics, and content analysis, which involves coding data into predetermined catego- ries. Qualitative analysis methods can be combined with quantitative procedures in mixed-method designs. Triangulation designs involve using both qualitative and quantitative analysis methods on the same data, while embedded designs give more weight to one method. Explan- atory designs incorporate qualitative analysis to explain quantitative results, while exploratory designs conduct quantitative analysis based on qualitative findings. The combination of qualitative and quantitative methods can be flexible depending on the research questions at hand.

# UNIT 6

## RESEARCH ETHICS

**STUDY GOALS**

On completion of this unit, you will be able to ...

* name some of the most well-known and controversial experiments in the field of psy- chology.
* comprehend and describe the importance of informed consent in research studies.
* explain concepts of anonymity and pseudo-anonymity in research.
* describe the concept of deception in research and understand the ethical considera- tions surrounding its permissibility.
* outline the key principles and guidelines of good scientific practice.

### 6. RESEARCH ETHICS

#### Introduction

Psychological research primarily focuses on studying human behavior, attitudes, thoughts, emotions, and brain function. Given that human participants are involved in such research, it is crucial to establish specific regulations to ensure ethical treatment (Goodwin, 2010). Ethics in research refers to a set of guidelines that govern the conduct of studies involving human participants. The primary objective is to safeguard the well-being of participants and prevent any harm during their involvement in research projects. It is the moral responsibility of research teams to protect human participants, which is why a set of principles and rules of conduct have been established (Goodwin, 2010; McLeod, 2015). The American Psychological Association (APA), established in 1892, publishes rules and regulations for research conduct (Pickren & Rutherford, 2018). Similarly, the British Psychological Society (BPS), founded in 1901, provides similar guidelines in the United Kingdom. In Germany, the German Society for Psychology (Deutsche Gesellschaf t für Psy- chologie [DGP]), established in 1904, also publishes rules for research conduct. While the rules rarely vary between countries, there is significant overlap between the regulations set forth by each society, and many of the regulations effective in Western societies are based on the guidelines established by the APA (Bunn, 2021).

Many universities and research institutions have established ethics committees comprised of experts from various disciplines who assess the ethical acceptability of research projects. Before initiating a research project, a research team must obtain ethical approval by submitting a research proposal to the ethics committee for evaluation. The proposal must adhere to specific guidelines provided by the committee and include all relevant project details. The committee then either grants approval or provides recommendations for necessary changes to ensure compliance with ethical standards. Ethical approval is also required for the subsequent publication of research findings. Most peer reviewed- sci- entific journals require the research team to provide a copy of the original ethical appro- val, without which the data cannot be published.

#### Well-Known Social Psychological Studies with Problems in Research Ethics

In psychology and social sciences, certain experiments have revolutionized our under- standing of human psychology. However, they have also received significant criticism due to ethical concerns about the conducting of experiments. Today, these experiments would not be permitted due to ethical considerations. These tests have greatly influenced the discussion on moral limits that should not be exceeded while conducting research. How- ever, to what extent should obtaining new knowledge justify the violation of ethical stand- ards?

##### The Little Albert Experiment

The experiment conducted by the psychologist Watson in 1920 focused on classical condi- tioning (Watson & Rayner, 1920). Classical conditioning is a type of learning in which behavior is shaped through association. The experiment used a nine-month-old baby, nicknamed Albert, as a volunteer and involved furry objects (Harris, 1979). In the experi- ment, Albert was initially given a white furry object to play with, and he showed joy and liking towards it. However, while Albert was playing with the furry object, Watson and his team would create a loud noise behind him to intentionally frighten him. This procedure was repeated multiple times. As a result, Albert was conditioned to develop fear responses not only to the specific furry object used in the experiment but also to other white furry objects (Watson & Rayner, 1920). It is important to note that by today’s ethical standards, this study would be considered unethical. Albert was exposed to distressing conditions, as fear responses were intentionally induced without his awareness or con- sent. Additionally, as a baby, he was unable to provide informed consent for participation in the study.

##### The Milgram Experiment

The Milgram experiment, conducted by Stanley Milgram in 1963, aimed to indirectly study human obedience. The experiment sought to understand the acts of extreme violence and genocide that occurred during World War II in Nazi Germany, where individuals followed orders from superior officers despite conflicting with their personal conscience. Milgram, a research psychologist at Yale University, designed and conducted the experiment (Mil- gram, 1963). In the experimental procedure, participants were paired with another individ- ual who they believed to be another participant but was actually a member of the research team or an actor. Participants were informed that they would be randomly assigned to one of two conditions: the learner or the teacher. Through deception, the real participant was always assigned the role of the teacher, while the member of the research team or actor played the role of the learner.



**DECEPTION**

In research, deception refers to the intentional misleading or partial disclosure of information to research participants regarding the purposes of the study. While deception should generally be avoided, there are rare occasions where it may be necessary as part of the experimental manipulation, as long as it does not cause harm or distress to the participants. It is crucial that at the conclusion of the experimental procedure, participants are provided with a full explanation of the procedures involved to ensure their understanding and to address any potential concerns or questions they may have.

The learner (actor or fake participant) was placed in a room alone and was instructed to learn word pairs while having electrodes attached to them. The teacher (the real partici- pant) would enter another room with the researcher and be in control of an electric shock

generator (which was actually fake, but the participant was unaware of this). The teacher’s role was to test the learner by reading one word and asking them to recall the other word of the pair. The instructions given to the teacher were to administer an electric shock to the learner whenever they gave a wrong answer. In reality, no electric current was involved, and the learner’s reactions were staged.

The experiment aimed to examine obedience to authority, and if the teacher hesitated to administer shocks, the experimenter would use authoritative language to encourage them to continue. Despite the participants’ awareness that strong electric shocks could be harmful or even fatal, 26 out of 40 (65%) participants continued to administer shocks up to the highest level of 450 volts, while 14 (35%) participants stopped the experiment early. Milgram interpreted these results as evidence that most people are willing to follow authoritarian instructions, even if it causes harm to others (Milgram, 1963). The Milgram experiment has indeed faced significant ethical criticism. Participants were subjected to intense internal pressure and were misled about the true purpose of the experiment. Their right to withdraw from the trial at any point, as they should have been able to, was vio- lated. Participants who expressed a desire to stop were pressured by the experimenter to continue. It is important to note that the exact same experimental procedure would not receive approval to be conducted today due to these ethical concerns.

##### The Stanford Prison Experiment

The experiment was conducted by Zimbardo at Stanford University USA in the early 1970s (Haney et al., 1973). Zimbardo aimed to investigate how individuals adapt to societal roles. He designed the experiment to explore whether the tension between prisoners and prison guards was primarily influenced by personality traits or the prison environment. Participants were randomly assigned the roles of either prisoners or guards, and a simu- lated prison was created in the basement of Stanford University where the prisoner partic- ipants were held. The guards were given specific shifts and instructed to treat the prison- ers as they would in a real-life prison setting. Both groups were aware that they were participating in a simulated role-playing experiment.

**Informed consent** This principle is funda- mental in all research involving human partici- pants. It entails providing each participant with suf- ficient information about the procedures and objectives of the research project before they agree

to participate.

It became evident to the research team early on in the experiment that participants in both the prisoner and guard roles quickly adapted to their assigned roles (Haney et al., 1973). However, the experiment had to be terminated earlier than planned, on the sixth day instead of the initially intended two weeks, due to escalating conflicts and the initia- tion of a rebellion by the prisoners as early as the second day. The guards displayed sadis- tic behavior, while the prisoners exhibited increased levels of anxiety and depression. The experiment’s findings concluded that individuals can readily adapt to social roles based on societal expectations (Haney et al., 1973). Ethical concerns surrounding this study pri- marily revolve around the lack of fully **informed consent**. Participants were not fully informed about what would transpire during the experiment or how the situation would progress, as even the experimenter himself was uncertain. Additionally, participants assigned to the prisoner group did not consent to their “arrest” at their homes, which was how their involvement in the study began (McLeod, 2020).

#### Ethical Handling of Subjects

Depending on their area of research investigation, such as clinical psychology, educational sciences, or anthropology, researchers must adhere to ethical guidelines, although priori- ties may vary. Nonetheless, the following ethical guidelines and regulations for dealing with human test participants are relevant to any area of psychology research.

##### Informed Consent

Research participation is based on voluntary consent, which is obtained after providing comprehensive information to potential participants. The research team provides a detailed description of the project, its aims, procedures, and potential risks involved (APA, 2017). Participants have the opportunity to ask questions and, once satisfied, can provide their informed consent by signing a declaration. This consent serves as a contract between the research team and participants, protecting the rights of both parties. Impor- tantly, participants are free to withdraw from the study at any point without providing a reason. Informed consent is only considered valid when participants have received suffi- cient information and had all their questions addressed (APA, 2017). Consent is necessary in research studies because they often involve the collection of personal information, genetic material, or other biometric data. It is an absolute requirement, as mandated by European law, for research studies conducted within the EU to adhere to the General Data Protection Regulations (GDPR) concerning data protection and privacy.

It is essential to ensure that participants have a clear understanding of the information presented in the study and are capable of making informed decisions about their partici- pation. Informed consent documents should be written in plain language that is easily understandable to participants. In research involving children or individuals with limited decision-making capacity, consent is obtained from their representatives, such as parents or legal guardians. However, children should also be informed about the study and have the opportunity to provide their verbal assent to participate. It is crucial that no child or person is compelled to participate in a study against their will, even if their legal guardian or parent has provided consent (APA, 2017).

There are certain exceptions where informed consent may not be required in research. For instance, in risk-free studies, it may be permissible to omit informed consent. Additionally, observing the behavior of individuals in public places or using publicly available docu- ments, including online material, such as on Instagram, can be conducted without the explicit consent of the individuals being observed. However, it is crucial to ensure that the data collection procedures do not cause any burden or harm to the individuals involved (Döring & Bortz, 2016).

##### Confidentiality and Anonymity

Respecting the confidentiality of research participants is of utmost importance. Research- ers must take necessary measures to ensure that all data and information collected remains confidential, with no possibility of identifying the individuals who participated in the study. Decisions should be made regarding the extent of personal data collection and

how it is handled (APA, 2017). The GDPR outlines regulations that apply to research and the handling of personal data within the EU. Personal data should be collected and stored anonymously, meaning that no conclusions can be drawn about the individuals being studied, and the collected data cannot be linked to any specific person (APA, 2017). In research, once a person enters a study, their name is replaced with a randomly generated ID consisting of letters and digits. Only these IDs are stored in the records, making it impossible to identify the individuals. The names of participants may still exist, and if indi- viduals have explicitly expressed their willingness to be contacted again, they can be con- tacted for future studies. However, it is crucial to ensure that no personal data or any other information collected as part of the research project can be linked to the names, thus ensuring anonymity (APA, 2017).

##### Deception and Debriefing

Deception is a practice that is sometimes employed in research studies or designs. For instance, in the Milgram experiment, participants were deceived into believing that the other participant involved was a genuine participant when in reality they were an actor. Deception in research can take two forms: either by withholding certain information about the study at the beginning or by providing participants with incomplete or inaccurate information regarding the study’s purpose or procedures (Goodwin, 2010). In general, deception should be avoided in research, unless there is a strong likelihood of gaining sig- nificant scientific or practical knowledge and no suitable alternative approach is available. Many experimental designs in psychology involve deception as it is deemed necessary for addressing specific research questions. However, in cases where deception is employed, it is the responsibility of the researcher to inform participants about the deception as soon as their participation is complete. This process is known as debriefing (Goodwin, 2010).

The Iowa Gambling Task (IGT) is a popular psychological test that simulates real-life deci- sion-making abilities in humans. Participants are presented with four decks of cards and are instructed to choose cards in order to earn money (represented by facsimile) (Bechara et al., 1994). They are informed that some decks are more advantageous than others in terms of winning and losing money. However, what participants are not told is that they need to figure out which decks are advantageous and which are disadvantageous. Partici- pants make card selections one at a time, aiming to maximize their earnings. After playing the IGT, participants are provided with a debriefing session where researchers explain the true nature of the task and how they should have played to maximize their winnings. The deception in this test is necessary to examine implicit memory abilities, as participants cannot be informed about the contingencies involved. Importantly, no participant experi- ences distress or discomfort during the test, and the debriefing ensures that they fully understand the purpose and procedures of the study.

##### Participation Reimbursement

Participation in research studies is based on voluntary consent, meaning that individuals are given the choice to decide whether or not they want to participate. No one is compel- led or forced to become a research participant unless they genuinely wish to do so. This principle of voluntary participation ensures that individuals have autonomy over their decision to take part in research (APA, 2017). Participants in research studies typically

receive payment as compensation for their time, effort, and any study-related expenses they may incur. This payment is intended to acknowledge and compensate participants for their contribution to the research. It is important to note that the payment provided is not considered a form of financial incentive or inducement, but rather a fair compensation for their involvement.

##### Risk-Benefit Ratio

Research involving human or animal subjects should not cause any physical or psycholog- ical harm. However, participation in a study may trigger negative emotions or thoughts in participants. The question remains whether researchers have ignored the risk of harm for the sake of gaining knowledge (i.e., benefit) and whether this constitutes an unethical practice.

Over the years, numerous research studies have centered on addiction, specifically alco- hol use disorders. These studies analyze the effects of alcohol on human behavior, thought processes, brain activity, and genetics. To investigate this impact, individuals who experience alcohol dependency are invited to take part in research studies. Participants in the study undergo various experimental tests, such as viewing images of alcoholic drinks and evaluating their propensity to have a drink after seeing a picture of a bottle of wine. They also smell alcoholic beverages and rate their sensations. It is a risky situation for these individuals, particularly since the study often involves inviting patients in remission from alcoholism. Participating in such a study could increase the likelihood of drinking again, posing a potential risk. However, ethical approval for these studies indicates that this risk is manageable. Moreover, the risk faced by patients in these studies does not exceed that of society at large, where alcoholic beverages are widely available in bars, supermarkets, and grocery stores. Therefore, participating in a study does not inherently expose individuals to the stimulus they seek to avoid (alcohol) and consequently increase their risk of drinking. Additionally, it is essential to note that research participation is vol- untary, and all participants receive complete information regarding the procedures. Thus, any patient or participant who experiences alcohol use disorders can choose whether or not to participate, particularly when risks are known to them.

Researchers and participants weigh the risks associated with the study to make decisions. It is important to guarantee a balance between the benefits and risks and avoid any irre- versible harm to both parties. Researchers may also face risks such as psychological stress from hearing distressing stories during interviews or potential danger en route to the sur- vey location. All necessary precautions should be taken to mitigate any potential risks, such as providing appropriate psychosocial support for the participants and ensuring reg- ular supervision of the interviewers.

##### Outline of Informed Consent

The information regarding the study should strive for maximum transparency, ensuring that all relevant aspects are provided to the participants. This includes offering clear explanations and comprehensive details about

* the purpose of the research, anticipated duration of the investigation, and the proce- dure;
* the participants’ right to refuse or terminate their participation, even if the study has already started;
* the option to prematurely terminate participation, without any negative consequences.
* foreseeable factors that may influence the willingness to participate (e.g., potential risks or concerns about the knowledge that may be gained through the research);
* the guarantee of confidentiality and anonymity, including any limitations that may apply and the measures taken to ensure them, if applicable;
* the storage, use, and deletion of the data. This includes information on where the data will be stored, who will have access to it, and how long the research team intends to retain it;
* the reimbursement of participation; as well as
* the contact information of responsible individuals within the research team who can address any questions or complaints.

#### Rules of Good Scientific Practice

Irrespective of whether individuals are subjects in a study or not, researchers across all disciplines adhere to the principles of good scientific practice (GSP). These principles offer guidance for a responsible approach to generating, interpreting, and publishing results (DFG, 2019). Some general rules related to GSP that should be applied to all research, not just psychological, are outlined here. GSP pertains to the scientific methods employed, which should be selected based on the current state of research (“state of the art”). Scien- tific findings should undergo critical evaluation through peer-review processes. Addition- ally, scientific data should be securely stored and retained for an agreed-upon period of time, and contributions from competitors or cooperation partners should be acknowl- edged (APA, 2017; DFG, 2019).

##### Training, Responsibility, Supervision

The responsibility for adhering to GSP and its standards lies with the principal investigator (PI) of the research project and their team. The head of the department where the research is conducted also shares accountability. The PI bears sole responsibility for supervising their research team and any students involved in the project. It is the PI’s duty to ensure that all individuals in the research team are appropriately trained, aware of their responsibilities, and capable of conducting research in accordance with ethical guidelines and regulations (APA, 2017; DFG, 2019).

##### Ethical Research Practice

Research involving human participants must uphold the dignity, rights, safety, and well- being of the individuals involved. Researchers are obligated to ensure absolute confiden- tiality and security of personal information pertaining to the research participants. The research project must also adhere to any legal requirements, such as those outlined by the EU GDPR (DFG, 2019).

##### Data Collection and Tools

Principal investigators and all individuals involved in a research project should establish clear guidelines from the outset regarding the ownership, usage, storage, and disposal of data and information collected during the research, as well as the dissemination of research results and the tools provided for data collection purposes (including best practi- ces for collecting high-quality data). Researchers are encouraged to consider the long- term maintenance of the project’s research data for use by future researchers (DFG, 2019).

##### Record Keeping

Researchers are required to maintain clear and accurate records of the procedures fol- lowed during the research process, including documentation of preliminary results obtained and final outcomes (DFG, 2019).

##### Honesty and Integrity

Researchers are expected to uphold honesty not only in their own research endeavors but also in their interactions with other researchers, encompassing overall research work, experimental design, data generation and analysis, funding applications, publication of results, and peer review of other researchers’ work. Researchers must refrain from engag- ing in plagiarism, deception, fabrication, or falsification of results, as these are considered serious disciplinary offenses (DFG, 2019).

##### Dissemination and Publication of Results

When publishing or disseminating research findings, researchers should prioritize the benefit to patients, the general public, and the scientific community (DFG, 2019).



**SUMMARY**

Ethical considerations hold particular significance in psychological research due to its focus on individuals and their subjective experiences and behaviors. Notable studies with ethical concerns, such as the Mil- gram experiment, have influenced discussions about the ethical treat- ment of research participants. Ethical principles and standards of good scientific practice are primarily established by professional organiza- tions that oversee research activities globally, such as the Association of German Psychologists (BDP), the British Psychological Society (BPS), and the American Psychological Association (APA). Ethical principles and guidelines for working with individuals should be integrated at all stages of the research process, from topic selection to result presentation and dissemination. The standard procedure involves presenting a research project to an ethics committee, which assesses the ethical justification for the proposed research. Key ethical considerations include a favora- ble risk-benefit ratio, obtaining informed consent, and anonymizing per-

sonal data. Furthermore, researchers are expected to adhere to the prin- ciples of good scientific practice, which guide a responsible approach to generating, interpreting, and publishing results, ensuring quality and defining scientific misconduct, such as plagiarism and falsification.